

PL-TR-92-2227

AD-A259 894



2

REGIONAL-SCALE ANALYSIS AND FORECASTING

John Burgeson
Christopher Redder
Isidore Halberstam
Ralph Shapiro

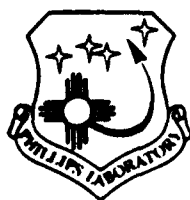
DTIC
ELECTE
DEC 16 1992
S A D

Science and Technology Corporation
P.O. Box 7390
Hampton, VA 23666

September 1992

Final Report
May 1991 to September 1992

Approved for public release;
Distribution unlimited



PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-5000

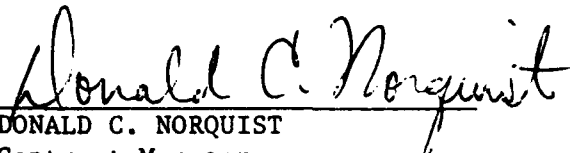
92-31471



180pgs

92 12 15 037

"This technical report has been reviewed and is approved for publication"


DONALD C. NORQUIST
Contract Manager


DONALD A. CHISHOLM
Branch Chief


ROBERT A. McCLATCHEY
Division Director

This report has been reviewed by the ESC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS)

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/TSI, Hanscom AFB, MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. Agency Use Only (Leave blank).		2. Report Date. September 1992	3. Report Type and Dates Covered. Final Report — May 1991-September 1992	
4. Title and Subtitle. Regional-Scale Analysis and Forecasting			5. Funding Numbers. PE 63707F PR 2688 TA 01 WU JC	
6. Author(s). John Burgeson Christopher Redder Isidore Halberstam* Ralph Shapiro*			Contract F19628-89-C-0167	
7. Performing Organization Name(s) and Address(es). Science and Technology Corporation 109 Massachusetts Avenue Lexington, MA 02173			8. Performing Organization Report Number. STC Technical Report 2660	
9. Sponsoring/Monitoring Agency Name(s) and Address(es). Phillips Laboratory Hanscom AFB, MA 01731-5000 Contract Manager: Donald Norquist/GPAP			10. Sponsoring/Monitoring Agency Report Number. PL-TR-92-2227	
11. Supplementary Notes. *Hughes STX Corporation				
12a. Distribution/Availability Statement. Approved for public release; distribution unlimited			12b. Distribution Code.	
13. Abstract (Maximum 200 words). This final report completes the description, begun in the interim report, of the development of the regional-scale analysis and forecasting procedure (RAP). The RAP is based on multivariate optimum interpolation of observations selected rigorously by the forward stepwise regression procedure. The complete observation set undergoes a unique buddy check that carefully eliminates those observations that adversely affected the analysis. The scheme used regionally-developed univariate and bivariate forecast error correlations, which were a function of distance between observations. The models are stratified by region, season, standard atmospheric levels, and forecast lengths. The models are superior to the model described in the interim report, and the models are shown to be applicable to rawinsonde and satellite observations. Results of verification experiments with numerical analyses of real data showed that the RAP concept is viable. Observing system simulation experiments (OSSEs) further verified the quality of RAP and the validity of the developed correlation models. In addition, the OSSEs provided relative confidence estimates of the quality of RAP analyses. An interesting forecasting procedure was recommended that extended the RAP analyses into a short-range forecast.				
14. Subject Terms. Optimum interpolation, stepwise regression, forecast error correlations, observation system simulation experiments, numerical weather analysis, numerical weather forecast models, and observation buddy checking			15. Number of Pages. 182	
			16. Price Code.	
17. Security Classification of Report. Unclassified	18. Security Classification of This Page. Unclassified	19. Security Classification of Abstract. Unclassified	20. Limitation of Abstract. SAR	

FOREWORD

Science and Technology Corporation (STC) is pleased to submit this final report entitled "Regional-Scale Analysis and Forecasting" as part of Contract No. F19628-89-C-0167. The objective was to develop and demonstrate a regional-scale analysis and forecasting procedure (RAP) using optimum interpolation in both real and simulation modes. Using RAP, all available meteorological data were assimilated and fused efficiently into a high resolution analysis of mass, motion, and moisture fields. This report briefly reviews the five major tasks required by this contract and discussed in greater detail in the RAP Interim Report. The final status of each of these tasks is fully described, specifically, the operational optimum interpolation scheme, the development of the forecast error correlation models, the verification of experiments in numerical analysis, the analysis of observation system simulation experiments, and the recommendation for extending the RAP analysis to a 12-hr forecast. Last, but not least, the valuable technical assistance provided by Mr. Donald Norquist, Contract Manager, is acknowledged and greatly appreciated.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution /	
Availability Codes	
Dist	Availability Codes
A-1	

DTIC QUALITY INSPECTED 2

ABSTRACT

This final report completes the description, begun in the interim report, of the development of the regional-scale analysis and forecasting procedure (RAP). The RAP is based on multivariate optimum interpolation of observations selected rigorously by the forward stepwise regression procedure. The complete observation set undergoes a unique buddy check that carefully eliminates those observations that adversely affected the analysis. The scheme used regionally-developed univariate and bivariate forecast error correlations, which were a function of distance between observations. The models are stratified by region, season, standard atmospheric levels, and forecast lengths. The models are superior to the model described in the interim report, and the models are shown to be applicable to rawinsonde and satellite observations. Results of verification experiments with numerical analyses of real data showed that the RAP concept is viable. Observing system simulation experiments (OSSEs) further verified the quality of RAP and the validity of the developed correlation models. In addition, the OSSEs provided relative confidence estimates of the quality of RAP analyses. An interesting forecasting procedure was recommended that extended the RAP analyses into a short-range forecast.

CONTENTS

	<u>Page</u>
FOREWORD	iii
ABSTRACT	v
LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF ACRONYMS	xv
 1. INTRODUCTION	 1
2. REGIONAL ANALYSIS PROCEDURE TASKS	4
2.1 STC TASK 1	4
2.1.1 A Practical Application of Two-Dimensional Multivariate Optimum Interpolation in the Development of an Analysis on Mandatory Levels ...	4
2.1.2 The Buddy Check	12
2.1.3 The Incorporation of Satellite Observations with Rawinsonde Observations	16
2.2 STC TASK 2	21
2.2.1 Modeling Observation Errors	21
2.2.2 Forecast Error Correlations	21
2.2.2.1 The structure of spatial forecast error correlations at the 500-mb level	21
2.2.2.2 Forecast error correlations at 850- and 150-mb levels	25
2.2.2.3 Univariate and bivariate vertical error correlations	25
2.2.2.4 Temporal forecast error correlations	34
2.2.2.5 Summary of forecast error correlation models	41
2.2.3 Forecast Error Correlations of Satellite Data	42
2.2.3.1 Non-normal distribution of forecast errors	43
2.2.3.2 Justification for using the forecast error correlation model developed for RAOB data to model satellite data	44
2.2.4 The Development of Forecast Error Correlation Models	49
2.2.5 Fitting Real Data to the Forecast Error Correlation Models in Practice ...	52
2.3 STC TASK 3	54
2.3.1 Verification of RAP Analyses	55
2.3.2 Comparison of RAP Analyses with HIRAS Analyses	59
2.4 STC TASK 4	64
2.4.1 Development of the OSSE Databases	66
2.4.2 An Analysis of the BASELINE OSSE Global Data Assimilation Fields ...	68
2.4.3 An Analysis of the CONTROL OSSE Global Data Assimilation Fields ...	86
2.4.4 An Analysis of the TACOBS OSSE Global Data Assimilation Fields ...	90

2.4.5	The RWM Forecasts from the OSSEs	96
2.4.6	Results of RAP Analyses from the OSSEs	100
2.5	STC TASK 5	115
2.5.1	Background	116
2.5.2	Extension of Solution Discussed in the Proposal	116
2.5.3	Literature Survey	117
2.5.4	Recommendation	120
3.	SUMMARY AND CONCLUSIONS	121
APPENDICES		
A.	Case Studies of the Buddy Check	A-1
B.	Results of Observing System Simulation Experiments at 850 and 300 mb	B-1

LIST OF FIGURES

		<u>Page</u>
Figure 1.	The RAP grid, marked by dots, superimposed upon the RWM region called Eurasia.	2
Figure 2.	The RAP grid, marked by dots, superimposed upon the RWM region called central America.	3
Figure 3.	Depiction of the order of observation selection by the FSR procedure with respect to a grid point at the origin.	11
Figure 4.	The $\Delta Z_i - \Delta Z_j$ error correlations along the north-south direction at 850 mb (8), 500 mb (5), and 150 mb (1), as a function of distance between observations. . . .	26
Figure 5.	The average $\Delta U_i - \Delta U_j$ correlations $[r_{ij}(\Delta U_i, \Delta U_j)]$ from quadrant 2 (or 4) at 850 mb (8), 500 mb (5), and 150 mb (1), as a function of the distance between observations points i and j.	27
Figure 6.	The average $\Delta Z_i - \Delta U_j$ correlations, $[r_{ij}(\Delta Z_i, \Delta U_j)]$, from quadrants 1 and 3 at 850 mb (8), 500 mb (5), and 150 mb (1), where the N or S following the 1, 5, or 8 identifies the direction (north or south) from observation point i to point j, as a function of the distance between observations	28
Figure 7.	Approximate frequency distribution of 500-mb height errors from bin 10 (450 km < separation distance < 501 km).	45
Figure 8.	Height-height forecast error correlations at 500 mb in the north-south direction modeled from RAOB data (r), from interpolated RAOB data (i), and from interpolated satellite data (s).	47
Figure 9.	A depiction of the process for generating the OSSE databases, from 16 January 1979 at 00 UTC, which is the 00 hr for the T-106 forecasts.	65
Figure 10a.	Simulated GDA 48-hr analysis of 500-mb heights (dm) and temperature (C) valid 21 January 1979, 00 UTC over Eurasia for BASELINE.	69
Figure 10b.	Same as Figure 10a except for CONTROL.	70
Figure 10c.	Same as Figure 10a except for TACOBS.	71
Figure 11a.	Simulated GSA 48-hr analysis of 500-mb u-wind (m/s) valid 21 January 1979, 00 UTC over Eurasia for BASELINE.	72
Figure 11b.	Same as Figure 11a except for CONTROL.	73
Figure 11c.	Same as Figure 11a except for TACOBS.	74

Figure 12a.	Simulated GDA 48-hr analysis of 500-mb heights (dm) and temperature (C) valid 21 January 1979, 00 UTC over central America for BASELINE.	75
Figure 12b.	Same as Figure 12a except for CONTROL.	76
Figure 12c.	Same as Figure 12a except for TACOBS.	77
Figure 13.	Same as Figure 1 with the addition of a heavy line surrounding the data-denied region of the TACOBS scenario in Eurasia.	91
Figure 14.	Same as Figure 2 with the addition of a heavy line surrounding the data-denied region of the TACOBS scenario in central America.	92

LIST OF TABLES

	<u>Page</u>
Table 1. The Root-Mean-Square Errors of a 500-mb Analysis of Geopotential Height (Z) and Temperature (T) in the Eurasian Region	18
Table 2. The Root-Mean-Square Errors of a 500-mb Analysis of Geopotential Height (Z) and Temperature (T) in the Central American Region	19
Table 3. The Effect of Restricting the Use of SATOBs on an Analysis at 67 Rawinsonde Stations	20
Table 4. The Structure of the Vertical Correlations of Forecasts Errors Averaged over All Rawinsonde Sites at the Given Standard Level (k) with the Levels (l) Above and Below It for Each Window (Central America and Eurasia), Month (January and July), and Forecast Hour (12, 18, 24, and 36)	30
Table 5. Model for Illustrating the Temporal Error Correlations Between Elements A and B on Standard Vertical Levels as a Function of $\Delta t = 6, 12, 18,$ and 24 hr	35
Table 6. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta Z, \Delta Z)$	36
Table 7. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta Z, \Delta T)$	37
Table 8. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta T, \Delta Z)$	38
Table 9. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta T, \Delta T)$	39
Table 10. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta U, \Delta U)$	40
Table 11a. Root-Mean-Square Differences from Withheld Radiosonde Values at 850 mb of the First Guess Forecast, an Univariate Analysis with a Simple Correlation Model, a Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis	57
Table 11b. Root-Mean-Square Differences from Withheld Radiosonde Values at 500 mb of the First-Guess Forecast, an Univariate Analysis with a Simple Correlation Model, an Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis	57
Table 11c. Root-Mean-Square Differences from Withheld Radiosonde Values at 300 mb of the First-Guess Forecast, an Univariate Analysis with a Simple Correlation Model, an Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis	58

Table 12a.	The Ratio of the Root-Mean-Square Errors of 850-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 850-mb Analyses	60
Table 12b.	The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 500-mb Analyses	60
Table 12c.	The Ratio of the Root-Mean-Square Errors of 300-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 300-mb Analyses	61
Table 13a.	The Ratio of the Root-Mean-Square Errors of 850-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 850-mb Analyses	61
Table 13b.	The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the RMS Errors of HIRAS 500-mb Analyses	62
Table 13c.	The Ratio of the Root-Mean Square Errors of 300-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 300-mb Analyses	62
Table 14.	The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of RAP 500-mb Analyses Derived Using 36-Hr Forecast Error Correlation Models	63
Table 15.	The Root-Mean-Square Differences at Approximately 500 mb from the Three BASELINE OSSEs (shown separated by a slant) that Began at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979	79
Table 16.	The Root-Mean-Square Differences at Approximately 500 mb from the Three CONTROL OSSEs (shown separated by a slant) that Began at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979	87
Table 17.	The Difference Between the BASELINE OSSEs, Which Assimilate "Perfect" Observations, and the CONTROL OSSEs, Which Assimilate Simulated Observations	89
Table 18.	The Root-Mean-Square Differences in the TACOBS OSSE Near 500 mb. Three OSSEs, Each Separated by a Slant, Begin at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979 (that is, T-106 hours 00/120/240) and Last for 120 hr	93
Table 19.	The Difference Between the TACOBS and CONTROL (TACOBS-CONTROL) OSSEs at Approximately 500 mb	94

Table 20.	Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included over Eurasia (all initial times at 00 UTC)	101
Table 21.	Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess over Eurasia Forecasts on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included over Eurasia (all initial times at 00 UTC)	102
Table 22.	Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	103
Table 23.	Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	104
Table 24.	Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	105
Table 25.	Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	106
Table 26.	Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	107
Table 27.	Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	108
Table 28.	Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	109
Table 29.	Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the	

	Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)	110
Table 30.	Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)	111
Table 31.	Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)	112
Table 32.	Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)	113
Table 33.	Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central America Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)	114
Table 34.	Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Not Included (all initial times at 00 UTC)	115

LIST OF ACRONYMS

AF	analogue forecasting
AFGL	Air Force Geophysics Laboratory (now called PL/GP)
AFGWC	Air Force Global Weather Central
AIREPS	aircraft reports
ASAP	AFGL Statistical Analysis Package
ASCII	American Standard Code for Information Interchange
CAM	central American region
ECMWF	European Center for Medium-Range Weather Forecasts
EUR	Eurasian region
FG	first guess
FSR	forward stepwise regression
FGGE	First GARP Global Program Experiment
GARP	Global Atmospheric Research Program
GDA	global data assimilation
GSM	Global Spectral Model
HIRAS	High Resolution Analysis System
NWP	numerical weather prediction
OI	optimum interpolation
OSSE	Observing System Simulation Experiment
PL	Phillips Laboratory
PL/GP	Phillips Laboratory Geophysics Directorate
PLSC	Phillips Laboratory Supercomputer Center
RAOB	rawinsonde observation
RAP	Regional-Scale Analysis and Forecasting Procedure
RWAM	Relocatable Window Analysis Model
RWM	Relocatable Window Forecast Model
RWFMVER	Relocatable Window Forecast Model Verification
SATOB	satellite observation
STC	Science and Technology Corporation
UGDF	uniform gridded data field
UTC	Universal Time Constant
VAS	VISSR Atmospheric Sounder
VISSR	Visible and Infrared Spin-Scan Radiometer

1. INTRODUCTION

The regional-scale analysis and forecasting procedure (RAP) is a multivariate, multisource analysis scheme for relocatable, regional-scale applications. Specifically, the analysis procedure incorporates optimum interpolation for the numerical analysis of standard meteorological elements. The first guess for the analysis is provided by the Relocatable Window Forecast Model (RWM) from the Air Force Global Weather Central (AFGWC). Carefully validated observations of varying availability in time and space are used in an optimum sense to produce the best possible depiction of the variables within the regional atmospheric volume. A stepwise regression scheme selects the observations, based upon their intercorrelations, that will be incorporated into the analysis. The analysis procedure uses meteorological fields from various sources on a regional scale analysis grid, called uniform gridded data fields. Figures 1 and 2 show the RAP grid superimposed upon the RWM window over the regions of Eurasia and "central America," respectively. Note that the region identified as central America (Fig. 2) extends from southeastern United States southward to the Equator.

This report describes the five major tasks, including a detailed description of the prototype observation selection and optimum interpolation procedures, case studies of experiments in numerical analysis, and the observing system simulation experiments (OSSEs). This final report is preceded by the RAP Interim Report (Burgeson, et al. 1991), which documented the potential effectiveness of RAP, a relatively unsophisticated version at the time of the interim report.

The objective of the RAP was to develop and demonstrate a regional analysis procedure using optimum interpolation in both real and simulation modes. Using RAP, all available meteorological data were assimilated and fused efficiently into a high resolution analysis of mass, motion, and moisture fields. Verification experiments with real data demonstrated the effectiveness of the RAP concept. In addition, the OSSEs established confidence levels for RAP analyses.

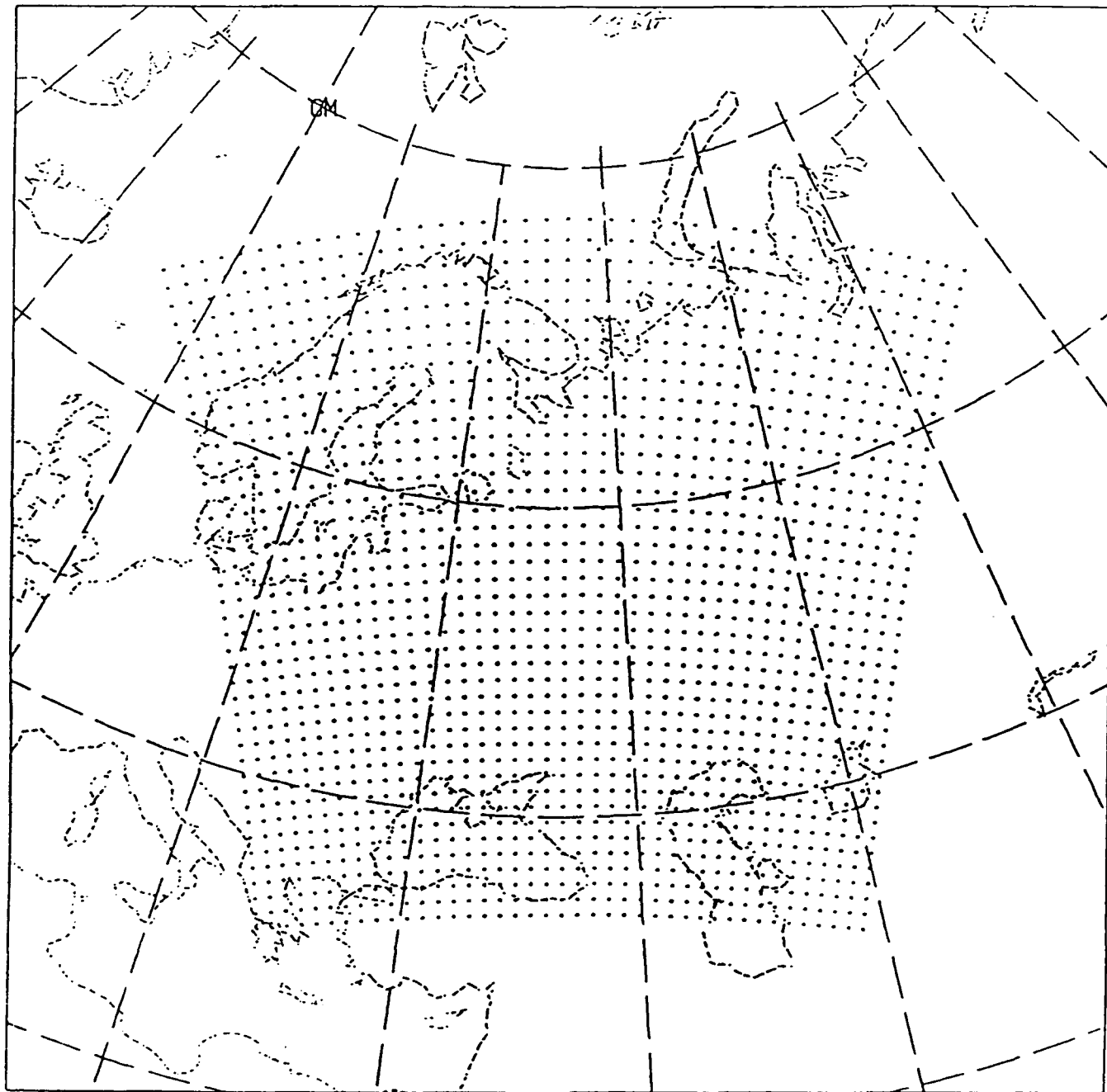


Figure 1. The RAP grid, marked by dots, superimposed upon the RWM region called Eurasia.

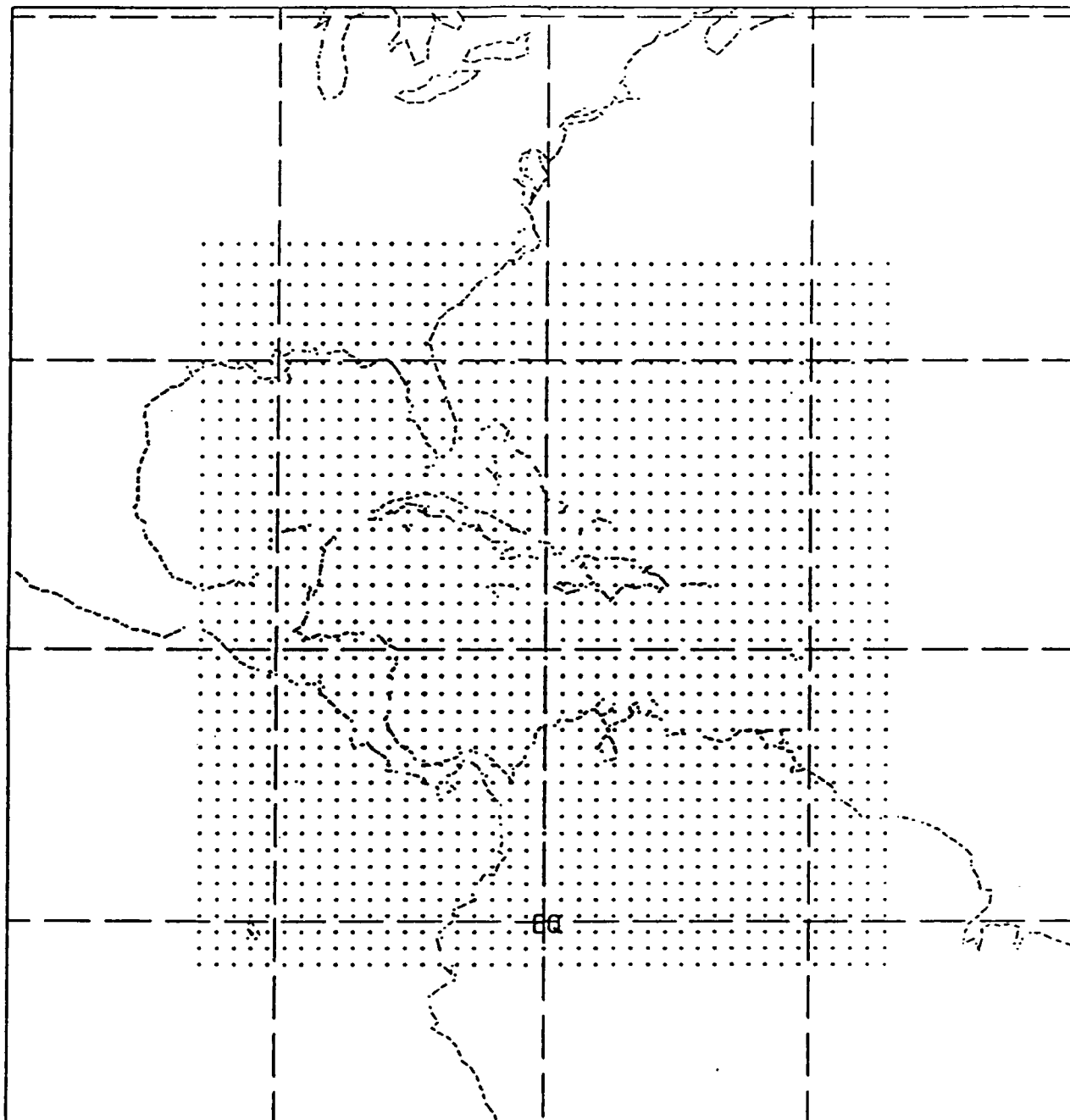


Figure 2. The RAP grid, marked by dots, superimposed upon the RWM region called central America.

2. REGIONAL ANALYSIS PROCEDURE TASKS

The research and development effort, which required a great deal of software modification and development, was specified in the Science and Technology Corporation (STC) Technical Report 3072, RAP Initial Work Plan, September 1989. Each of the five dependent tasks and the results of the work to complete them are described in detail.

2.1 STC TASK 1

Task 1 was the development of an automated relocatable, regional multivariate objective analysis procedure using optimum interpolation (OI). Two major subtasks were originally envisioned: (1) design and develop an OI scheme, and (2) incorporate the RAP analysis algorithm into the relocatability and variable resolution framework of the Relocatable Window Analysis Model (RWAM). In practice, however, all effort was directed to the design, development, and testing of the OI scheme, including observation selection; the incorporation of the RAP analysis algorithm into the relocatability and variable resolution framework of the Air Force Global Weather Central (AFGWC) RWAM¹ was cancelled.

The OI scheme, based on a multivariate, objective data selection by forward stepwise regression (FSR), has been carefully evaluated on all meteorological variables obtained from surface, rawinsonde and satellite observations. The variables were sea level pressure; the height of isobaric levels and the temperatures, both of which were available as rawinsonde and satellite data; and humidity, and u- and v-wind components, all of which were from rawinsonde data.

2.1.1 A Practical Application of Two-Dimensional Multivariate Optimum Interpolation in the Development of an Analysis on Mandatory Levels

This section briefly reviews and then greatly expands the discussion on optimum interpolation from the RAP Interim Report. Equation 1 describes how to calculate objectively an analyzed value at a grid point, g , in a preselected region containing N observation points.

¹The RWAM was not implemented at AFGWC; however, the RAP software is flexible so that a user can readily set the desired location and resolution of a window (region). Since the RWAM code was unavailable, no effort could be devoted to the second part of STC Task 1.

$$f_g^A = f_g^P + \sum_{i=1}^{N'} W_i (f_i^P - f_i^P) \quad (1)$$

where f_g^A is the estimated (analyzed) value of the variable at grid point g , f_g^P is a preliminary (forecast) value available only at grid points of the prediction model, f_g^P is a first guess forecast of the variable at g , f_i^P and f_i^P are the observed and first guess values of the variable at the observation points, respectively, $N' \leq N$ implies that it may be desirable to use less than all of the N available observations, and W_i is the weight assigned to the observation point i ($i = 1, 2, \dots, N'$). Note that $f_i^P - f_i^P$ can be written as Δ_i , the forecast error of f at i .

Given the weights, Eq. 1 can be readily solved. The weights can be determined from the basic equation of optimum interpolation as formulated by Keegan and Shapiro (1985):

$$\sum_{i=1}^{N'} W_i r(\Delta_i, \Delta_j) \sigma_{\Delta_i} \sigma_{\Delta_j} = r(\Delta_j, \Delta_g) \sigma_{\Delta_j} \sigma_{\Delta_g} + r(\Delta_j, e_g^P) \sigma_{\Delta_j} \sigma_{e_g^P} \quad (2)$$

where $j = 1, 2, \dots, N'$ available observation points; $r(\Delta_i, \Delta_j)$ is the forecast error correlation between the pair of observations at points i and j ; Δ_j is the forecast error of f at j , and Δ_g is the forecast error of f at grid point g ; σ_{Δ_i} , σ_{Δ_j} , and σ_{Δ_g} are the standard deviations of Δ at the i and j observation points and grid point g , respectively; $\Delta_i = f_i^P - f_i^P$ is the deviation of the observed value of the variable at an observation point i from the first-guess forecast (commonly referred to as the forecast error) interpolated to the point i ; and $e_g^P = f_g^P - f_g$ is the observation error at the grid point, where f_g is the observed value of a variable interpolated to the grid point. Since f_g is the true value of f at g and as such is an unknown quantity everywhere, e_g^P is also unknown. This is not a problem, however, for the OI procedure if the second term on the right side of Eq. 2 can be neglected (as will be discussed below), that is, if

$$r(\Delta_j, \Delta_g) \sigma_{\Delta_j} \sigma_{\Delta_g} \gg r(\Delta_j, e_g^P) \sigma_{\Delta_j} \sigma_{e_g^P} \quad (3)$$

Consider the variability of forecast errors at grid points σ_{Δ_g} , which should be much greater than the variability of observation errors at grid points. Forecasts usually verify well at some locations and poorly at others, especially in regions of active weather; however, observation errors are unlikely to differ much within a region the size of the RAP grid. Thus,

$$\sigma_{\Delta_i} \gg \sigma_{\epsilon_i} \quad (4)$$

is true in general. Furthermore, the correlation of forecast errors at observation points with forecast errors at grid points (especially close grid points) must be much greater than the correlation of forecast errors at observation points with observation errors at grid points. (There is no reason for them to be related except in the isolated case of an observing station that persistently reports erroneously.) Finally, $-1 \leq r \leq 1$ for all correlations, so Eq. 4 assures that Eq. 3 is generally satisfied. In addition, Bergman (1979), for example, noted that conventional observations are usually made with no knowledge of the forecast error; therefore, $r(\Delta_j, \epsilon_i^o)$ must be small. He argued further that the errors of observations from differing observation sensors are uncorrelated, and similarly, that errors of measuring two different meteorological variables, even if measured by the same instrument, are unlikely to be correlated. His discussion concluded with the exception of satellite observations, which produce smooth fields, and so may have errors that correlate with errors of forecast fields, which are also smooth.

Consequently, the inequality expressed by Eq. 4 is undoubtedly true for all observations except possibly satellite data. Section 2.2.3, however, shows that (at least for the relatively small data sample used) forecast error correlations for satellite data are similar to rawinsonde data. Accordingly, the error term in Eq. 2 probably can be safely ignored even when satellite data are part of the analysis.

Thus, neglecting the second term on the right side of Eq. 2 leads to yield

$$\sum_{i=1}^{N'} W_i \frac{\sigma_{\Delta_i}}{\sigma_{\Delta_g}} r(\Delta_i, \Delta_g) = r(\Delta_g, \Delta_g) \quad (5)$$

which can be expanded into a system of linear equations valid for grid point g and N' surrounding observations points. Let $r_{ij} = r(\Delta_i, \Delta_j)$, then Eq. 5 can be rewritten as

$$\begin{aligned} w'_1 r_{11} + w'_2 r_{12} + \dots + w'_n r_{1n} &= r_{1g} \\ w'_1 r_{21} + w'_2 r_{22} + \dots + w'_n r_{2n} &= r_{2g} \\ \vdots & \\ w'_1 r_{n1} + w'_2 r_{n2} + \dots + w'_n r_{nn} &= r_{ng} \end{aligned} \quad (6)$$

where w'_i is a normalized weight given by

$$w'_i = w_i \frac{\sigma_{\Delta_1}}{\sigma_{\Delta_i}} \quad (7)$$

Note that $r_{ij} = r_{ji}$ and $r_{ig} = r(\Delta_i, \Delta_g)$ is the correlation between Δ_i and Δ_g . Since the correlations and standard deviations can be determined, Eqs. 6 and 7 can be solved for W_i .

It is possible but impractical and unnecessary to solve the above system of equations at each and every grid point using all available observations in the region to determine an analysis. Even if using all available observations were practical, Sections 3.2, 3.3 and 5 of the RAP Interim Report showed that it is unnecessary. Optimum interpolation performs an equivalent analysis at a point whether using a small number of observations carefully selected by the FSR procedure, or using all the available observations within a given distance of the point. The RAP methodology of employing OI is described below in greater detail than in the interim report.

There are several possible approaches to calculating weights and selecting observations. The most straightforward and efficient one in this case is an indirect approach. The following analysis shows that the multiple correlation coefficient $r_{g,1,2,\dots,n}$, which is required to apply stepwise regression to the observation selection process, can be calculated from the weights.

Define the determinant D of the augmented correlation matrix from Eq. 6:

$$D = \begin{vmatrix} r_{gg} & r_{g1} & r_{g2} & \dots & r_{gn} \\ r_{1g} & r_{11} & r_{12} & \dots & r_{1n} \\ r_{2g} & r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{ng} & r_{n1} & r_{n2} & \dots & r_{nn} \end{vmatrix} \quad (8)$$

By Cramer's rule the weights can then be expressed as

$$w_i = \left(\frac{\sigma_{\Delta_i}}{\sigma_{\Delta_g}} \right) \frac{R_{gi}}{R_{gg}} \quad (9)$$

where R_{gg} is the minor of r_{gg} in D , R_{gi} is the product of $(-1)^{i+1}$ and the minor of r_{gi} in D , and σ_{Δ_i} and σ_{Δ_g} are the standard deviations of the forecast error of the meteorological elements at observation point i and grid point g , respectively.

Finally, $r_{g1,2,\dots,n}$ can be related to the weights by

$$r_{g1,2,\dots,n} = \sqrt{1 - \frac{D}{R_{gg}}} \quad (10)$$

From the law of minors in determinants (that is, a cofactor expansion along the top row in D), D can be expressed as

$$D = r_{gg}R_{gg} + r_{g1}R_{g1} + r_{g2}R_{g2} + \dots + r_{gn}R_{gn} \quad (11)$$

Substituting among Eqs. 9, 10, and 11 yields an expression relating the weights and the multiple correlation coefficient

$$r_{g1,2,\dots,n}^2 = 1 - r_{gg} + \sum_{i=1}^n r_{gi} w_i \left(\frac{\sigma_{\Delta_i}}{\sigma_{\Delta_g}} \right) \quad (12)$$

Note that the standard deviation of forecast errors is required at both observation points, where it can be readily calculated, and at grid points, where it is unknown. In practice, however, σ_{Δ_i} is calculated for each meteorological element on all levels for each season and region at observation points, and the average over all observations ($i = 1, 2, \dots, N$) is calculated for each level. These average values of σ_{Δ_i} are assumed to be the average values of σ_{Δ_g} too; they are used in Eqs. 6 and 12. (Thus, in an univariate analysis the ratio of the standard deviations $\sigma_{\Delta_i}/\sigma_{\Delta_g}$ is 1.)

The above discussion can help describe how an OI analysis at a given point is prepared. The point can be either a grid or an observation, but now assume a grid point, g . First, the weights are calculated from Eq. 6 for a relatively large preselected set of observations surrounding g . Then the multiple correlation coefficient, used in the FSR scheme to select the observations that will ultimately be used to

determine the optimum analysis at g , is calculated from the weights with Eq. 12. The procedure is completely general; in theory it applies equally to all dimensions. (See Section 2.2.2.5 for a brief discussion.) Due to time limitations in practice, however, only the horizontal correlations were used.

The FSR process of observation selection, described more formally in Section 3.3 of the RAP Interim Report, is summarized below.

1. Given a region containing a set of N quality-controlled observations, each typically having five variables (geopotential height, temperature, dewpoint, and zonal and meridional wind components), a subset of the 10 (or more) observations closest to g is established.
2. The most highly correlated observation, say A , is selected from this set first. Observation A is usually an univariate selection, that is, the first observation chosen is usually the same type of meteorological variable for which the analysis is being performed. However, other than correlation there is no restriction at any time regarding the choice of what meteorological variable will be chosen for an analysis (except for satellite observations, which are artificially restricted as discussed in Section 2.1.3). In practice, the procedure considers all 5 variables at each of the 10 observation points.
3. The multiple correlation coefficients $r_{g,A,i}$ are calculated between all pairs of remaining observations, that is, those pairs consisting of observation A and the other observations $i = 1, 2, \dots, N_s-1$, where N_s is the number of observation points in the subset. N_s is typically 10 points \times 5 variables = 50.
4. The next observation selected from the subset yields the largest value of all $r_{g,A,i}$ for the remaining observations. That is, observation B provides $r_{g,A,B} = \text{MAX} \{r_{g,A,i}\}$ for all remaining observations $i = 1, 2, \dots, N_s-1$.
5. Similarly, observation C is selected that maximizes $\{r_{g,A,B,i}\}$ for all remaining observations $i = 1, 2, \dots, N_s-2$.

6. The FSR procedure continues to take observations from the subset as long as the multiple correlation of the selected set of observations increases by a predetermined amount, as described below. In formal terms, the FSR procedure stops and rejects the last selected observation upon satisfaction of the criteria,

$$(r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_{k+1}}^2 - r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_k}^2) < 0.01 \quad (13)$$

where Δ_g is the forecast error at the grid point g , Δ_1 is the forecast error at observation point 1, etc., and Δ_k is the forecast error at observation point k .

However, implementing the FSR selection procedure with empirically fitted forecast error correlation models initially produced erratic weights and unrealistic multiple correlation coefficients and analysis values. Careful analysis of Eq. 6 showed gross physical inconsistencies in the intercorrelations and the correlations with analysis points. Eliminating these inconsistencies improved the results dramatically, but the procedure continued to generate unacceptable weights and, hence, analysis values at some of the points. A detailed inspection of the results after each observation selection by the FSR procedure revealed that the weights became erratic when the rate of increase in the multiple correlation coefficient accelerated. Based on experience with Eq. 6, slight inconsistencies in the correlation models may have accounted for the erratic weights. The effects of the inconsistencies do not become apparent until after the FSR procedure selected a few observations and solved larger matrix equations, whose weights may have been more sensitive to the small inconsistencies.

Constraining the FSR procedure to stop and reject the last selected observation, if the relationship

$$(r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_{k+1}}^2 - r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_k}^2) > (r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_k}^2 - r_{\Delta_g, \Delta_1, \Delta_2, \dots, \Delta_{k-1}}^2) \quad (14)$$

is true, eliminated the difficulties associated with the erratic weights. Inspection of the residuals showed that the FSR procedure stopped as a result of satisfying Eq. 14 for less than 10% of the analysis values. For these values the minimum multiple correlation coefficient residual was usually between 0.01 and 0.03, and in very few cases did the residual exceed 0.10.

Figure 3 shows an example of the order (that is, first A, then B, etc.) in which real observations, denoted by "o"s in their approximate relative positions, were selected. Note that the selection process

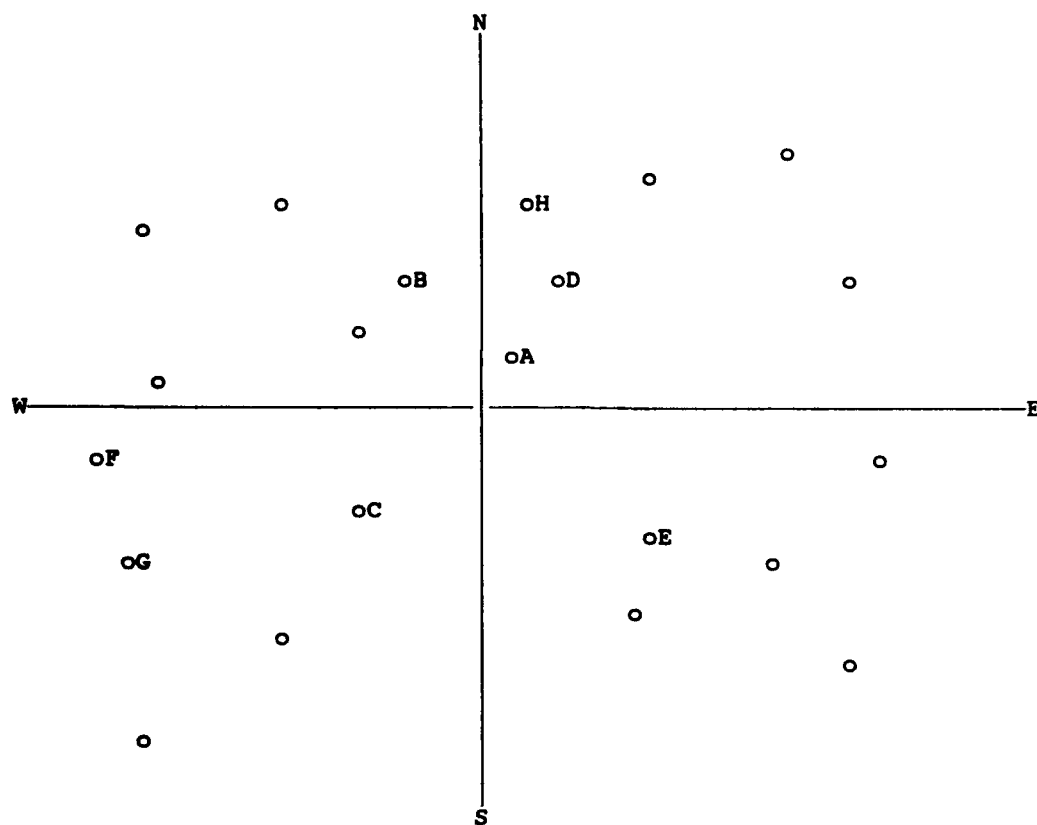


Figure 3. Depiction of the order of observation selection by the FSR procedure with respect to a grid point at the origin.

does not necessarily select the closest observations but rather those that provide the most information, as determined by the intercorrelations among all the observations, to the analysis at a point.

If N' is the total number of observations selected by the FSR procedure, then these N' observations are the only ones allowed to influence the analysis at g . The weights are calculated from the modeled error correlations associated with these observations using Eq. 6. The analyzed value of a variable at g is then calculated from Eq. 1. The entire procedure is repeated for all grid points until the analysis is completed.

2.1.2 The Buddy Check

The buddy check has undergone extensive revision and improvement since the previous version described in the interim report. One of the more significant changes is the derivation and use of an expression for the least-mean-square (lms) analysis error, derived from Eq. 1. By applying the method of assigning weights as formulated by Keegan and Shapiro (1985), the expression for the lms analysis error becomes

$$e_g^2 = \overline{(\Delta_g + e_g^o)^2} - \sigma_{\Delta_g}^2 \sum_{i=1}^N W_i' [r(\Delta_i, \Delta_g) + r(\Delta_i, e_g^o) \frac{\sigma_{\Delta_i}}{\sigma_{e_g^o}}] \quad (15)$$

where e_g is the root lms analysis error at point g , g refers to an analysis point that may be a grid point or an observation point, and the overbar represents an ensemble mean. In addition, recalling from Section 2.1.1, Δ_g and Δ_i are the deviations of the observation from the first guess (that is, the forecast residual) at the analysis point g and the observation point i , respectively; e_g^o is the observation error at point g ; σ_{Δ_g} , σ_{Δ_i} , and $\sigma_{e_g^o}$ are the standard deviations of Δ_g , Δ_i , and e_g^o , respectively; W_i is the weight given to observation point i ; W_i' is given by Eq. 7; $r(\Delta_i, \Delta_g)$ is the correlation between Δ_i and Δ_g ; and $r(\Delta_i, e_g^o)$ is the correlation between Δ_i and e_g^o .

Assuming the forecast bias and the mean observation error to be sufficiently small leads to the results,

$$\overline{\Delta_g} \ll [\overline{\Delta_g^2}]^{1/2} \rightarrow [\overline{\Delta_g^2}]^{1/2} = \sigma_{\Delta_g} \quad (16)$$

and

$$\overline{e_g^o} \ll [(\overline{e_g^{o2}})]^{1/2} \rightarrow [(\overline{e_g^{o2}})]^{1/2} = \sigma_{e_g^o} \quad (17)$$

In addition, if the correlations and covariances between observation errors and forecast residuals are negligible, then

$$\overline{\Delta_g e_g^o} \ll \overline{\Delta_g^2}, \quad \overline{\Delta_g e_g^o} \ll \overline{(e_g^o)^2} \quad (18)$$

and $r(\Delta_i, e_g^o) \ll r(\Delta_i, \Delta_g)$. With these assumptions the formula for the lms analysis error becomes

$$e_g^2 = \sigma_{\Delta_g}^2 \left[\frac{\sigma_{\epsilon_g}^2 + \sigma_{\Delta_g}^2}{\sigma_{\Delta_g}^2} - \sum_{i=1}^N W'_i r(\Delta_i, \Delta_g) \right] \quad (19)$$

According to Thiebaux et al. (1986), r_0 , the autocorrelation between forecast residuals (errors) at the same point, is given by

$$r_0 = \frac{\sigma_{\Delta_g}^2}{\sigma_{\Delta_g}^2 + \sigma_{\epsilon_g}^2} \quad (20)$$

and, therefore the expression for the root lms analysis error reduces to

$$e_g = \sigma_{\Delta_g} \sqrt{\frac{1}{r_0} - \sum_{i=1}^N W'_i r(\Delta_i, \Delta_g)} \quad (21)$$

where the first term under the square root has a minimum value of one and the second term, identified by Eq. 12 to be the multiple correlation coefficient with $r_{gg} = 1$, has a maximum value of one. Note that the lms analysis error depends on neither the first guess nor the observations values at points surrounding g . Instead, the location of the surrounding points partially determines the weights W'_i and, therefore, the root lms analysis error e_g .

The buddy check scheme uses the root lms analysis error e_g to gauge consistency of each observation with surrounding values by first calculating the error ratio $E_g = \epsilon_g^\wedge / e_g$, where $\epsilon_g^\wedge = f_g^\wedge - f_g^\circ$ is the analysis residual at point g , f_g° is the observed value and f_g^\wedge is the estimated value obtained by univariate analysis using Eq. 1 and excluding the observation at point g . The analysis f_g^\wedge and the analysis residual ϵ_g^\wedge , unlike the root lms analysis error e_g , depend on the first guess and observation values at points surrounding g . In a domain free of observation errors, both the numerator (i.e., analysis residual, ϵ_g^\wedge) and denominator (i.e., the root lms analysis error e_g) in the error ratio will be small whenever the corresponding analysis f_g^\wedge is good. Similarly, if the analysis is poor, both the numerator ϵ_g^\wedge and denominator e_g should be large keeping E_g in check. However, if E_g is unduly large, then either the observation f_g° has a large error or the analysis f_g^\wedge is under the influence of surrounding observations with large errors. Thus, E_g is useful for identifying observations as valid or candidates for rejection.

After calculating E_s at all observation points throughout the given domain, the buddy check selects a subset of observations and assumes it is valid (that is, the subset containing observations is assumed to have sufficiently small errors). This subset, called V (valid), is the baseline with which to test for consistency and validity of observations identified as candidates for rejection. Since observations with the smallest E_s are likely to have relatively small observation errors, an appropriate method for selecting members for subset V would be one that chooses observations with low E_s . In the buddy check scheme, all observations with E_s less than E_{tol} , the root-mean-square (rms) of E_s for the domain, become members of subset V.

The remaining observations, which have larger error ratios, are candidates for rejection and belong to another subset, R (rejected). Many observations, however, should not belong to this subset because these observation points have a large $|E_s|$ due to a large analysis residual, ϵ_i^A , primarily influenced by nearby observations with large errors. To test for these values, the scheme first sorts all members of subset R in ascending order of $|E_s|$ and, beginning with the first observation of the sorted subset, recalculates $|E_s|$ ($= |E'_s|$) without using any observations within the subset R. If $|E'_s| \leq E_{tol}$, then the test shows that the observation is consistent with neighboring observations in subset V. The scheme, therefore, transfers the observation from subset R to V and makes it eligible for use in subsequent analyses during this test.

Finally, the scheme tests the impact of each remaining observation in subset R on surrounding values by first sorting all remaining members in subset R in ascending order of $|E'_s|$ and calculating the constant C_s from the expression (in sorted order)

$$C_s = \frac{|E'_s|}{a E_{tol}} \sum_{i=1}^n (\epsilon_i^A - \bar{\epsilon}^A) - \sum_{i=1}^n \frac{e_i}{e'_i} (\epsilon_i^{A'} - \bar{\epsilon}'^A) \quad (22)$$

where n is the number of analysis points in subset V affected by the value being scrutinized (if $n = 0$, then the scheme uses the alternate expression $C_s = |E'_s| - a E_{tol}$), ϵ_i^A is the analyzed residual at point i calculated by including observations in subset R as well as subset V, $\bar{\epsilon}^A$ is the average residual of all observations in the domain, $\epsilon_i^{A'}$ is the analyzed residual calculated by excluding all observations in subset R, $\bar{\epsilon}'^A$ is the average analysis residual calculated by excluding all observations in subset R and held constant throughout this test (i.e., the scheme does not update $\bar{\epsilon}'^A$ when modifying subset V), e_i is the root rms analysis error calculated by including observations in subset R as well as subset V, e'_i is the root

lms analysis error calculated by excluding any observation in subset R, and a is an error tolerance factor ($= 3.0$). If C_g is negative, the inclusion of the observation produces an analysis acceptable enough to warrant transfer to subset V and preserve the value. The observation is then eligible for use in subsequent analysis for this test. If C_g is positive, then the removal of the observation results in analysis errors at the influenced points low enough to justify rejection of the observation. If the observation is a wind component, then the scheme rejects the other component as well since both u- and v-winds depend on the same measured quantities, speed, and direction.

By including the weight $[|E'_g|/(aE_{tol})]$, the test considers the consistency of the previous analysis at the point being scrutinized with surrounding observations. This statement becomes apparent by noting that if $[|E'_g|/(aE_{tol})] < 1$ and $e_i = e'_i$ for all i , then including the observed value must produce an inferior analysis before the scheme will reject the value. Similarly, $[|E'_g|/(aE_{tol})] > 1$ and $e_i = e'_i$ for all i implies that the analysis must improve before the scheme preserves the value. The factor e_i/e'_i accounts for the expected changes in the analysis residual due to withholding observations in subset R from the analysis.

Prior to implementing the buddy check, the observations undergo a gross error check. A gross error check scheme rejects any observation whose forecast residual exceeds aF_{tol} . The variable F_{tol} is the forecast residual tolerance determined by implementing the same procedure as that used to calculate E_{tol} except that the method considers forecast residuals instead of error ratios. In addition, the scheme rejects both u- and v- wind components if either component fails the gross check.

Tables 1-10 of Appendix A illustrate the effect of both the buddy and gross error checks on the analysis using a 12 hr first guess forecast and rawinsonde observations valid 1 January 1989, 12 UTC over Eurasia and central America. An inspection of the tables shows that the buddy and gross error checks significantly reduced the root-mean-square of the analysis residuals at point g, ϵ_g^A (or errors). In some cases, the root-mean-square of ϵ_g^A initially exceeded but finally fell to a value significantly less than the root-mean-square of the forecast residual.

Table 11 of Appendix A contains a summary of the results in Tables 1-10 and consists of the percentage of the observations removed and the reduction of the root-mean-square of ϵ_g^A . The expression for the percent of removal or reduction of a value is

$$\frac{X_i - X_f}{X_i} \times 100\% \quad (23)$$

where X_i and X_f are the initial and final values, respectively, and the value X in this study is either the number of observations or the root-mean-square of ϵ_i^A . The percentages are averages over the pressure levels 850 mb, 500 mb, 300 mb, 150 mb, and 70 mb. Note the higher percentage rate of observation removal among wind values than scalar values, which is a result of rejecting both wind components even if only one component fails either the buddy or gross error check.

2.1.3 The Incorporation of Satellite Observations with Rawinsonde Observations

The procedure for selecting observations makes no assumption about their reliability; consequently, the OI procedure by itself chooses an observation—rawinsonde or satellite or aircraft—for an analysis at a point based only on the forecast error correlation models (discussed in Section 2.2) and the stepwise regression selection scheme. The models are identical for all observing systems (Section 2.2.3); therefore, the OI procedure in principle could be overwhelmed by satellite observations (SATOBS), which are available in much greater quantity but are generally perceived to be of lower quality than rawinsonde observations (RAOBs). The following experiment determined that satellite data must be discriminately incorporated into a RAP analysis and suggested how to use the SATOBs most effectively.

Appropriate preselection procedures can ensure that higher quality rawinsonde observations, which are the baseline for comparison and verification, would be reasonably chosen before lower quality satellite observations, even if they were closer and more numerous than the rawinsonde observations. The preselected set of observations then becomes the universe from which observations are selected to influence an analysis. The experiments were limited in scope with respect to using SATOBs in the OSSEs and were limited both in breadth and depth by the tight time constraint. The preselection criterion, developed here and used in the OSSEs, could likely be improved by the knowledge gained from more extensive research.

Several 500-mb analyses of geopotential height and temperature based on both rawinsonde and satellite observations were performed. The most meaningful experiments were over the central American region (Fig. 1), where satellite data had more impact than in the Eurasian region (Fig. 2), which has

excellent coverage of rawinsonde. The control considered here used only RAOBs to perform an analysis at rawinsonde stations (whose observations were not used for the analysis at that point). An uncontrolled experiment used both rawinsonde and satellite data together without any restrictions to perform an analysis at the rawinsonde stations.

Finally, other numerical analysis experiments used both rawinsonde and satellite data together but with restrictions placed on accepting SATOBs based on the proximity of rawinsonde stations. Prior to the preselection process of choosing those observations that were allowed to affect an analysis at a point, however, all satellite observations were scanned and any within a specified distance of a valid rawinsonde observation were flagged. The restriction in practice eliminates all SATOBs (of the height of geopotential surfaces and their temperature) that have a validated RAOB within a specified (250, 500, or 750 km) distance.

As shown in Table 1, which describes the experiments in the Eurasian region, the differences between the control (restricted use of SATOBs) and uncontrolled (no restrictions placed on the use of SATOBs) experiments are small. The control analysis had 500-mb height and temperature errors (all errors are root-mean-square, calculated by subtracting the analyzed value at a rawinsonde station from the value observed there) reduced by 70% and 51%, respectively, of the first guess errors. The uncontrolled analysis had 500-mb height errors reduced by 64%, and the 500-mb temperature errors reduced by 50% of the first guess errors. Eliminating SATOBs within 500 km of a RAOB had the desired effect of "improving" (goodness is determined by reducing the rms errors) the analysis slightly. Since rawinsonde stations are relatively close together, the 500-km restriction assures that SATOBs cannot overwhelm RAOBs in the Eurasian region. Note that the restriction provides experimental results nearly the same as the control experiment. Nevertheless, the results of the uncontrolled experiment suggest that unrestricted use of SATOBs, that is, using all satellite data regardless of how close to an analysis point, is at least slightly detrimental to the analysis.

Table 1. The Root-Mean-Square Errors of a 500-mb Analysis of Geopotential Height (Z) and Temperature (T) in the Eurasian Region

Element	First Guess	Control Analysis	Uncontrolled Analysis	Restricted 250 km	Restricted 500 km
Z	73.0	21.2	26.5	23.9	20.4
T	2.88	1.42	1.65	1.45	1.43

Table 2, which describes the experiments in the central American region, shows much larger differences between the controlled and uncontrolled experiments. Obviously, the uncontrolled (unrestricted) use of satellite data ruined the analysis, making it worse than the first guess. In the control analysis both the first guess 500-mb height and temperature errors were reduced by 51%. In the uncontrolled analysis, however, the 500-mb height errors were greater than the first guess errors, and the 500-mb temperature errors were the same. Clearly, satellite data overwhelmed the 67 rawinsonde stations, resulting in an unacceptably poor analysis. This result confirmed the assertion that satellite data should not be allowed to influence an analysis indiscriminately in regions where the RAOBs would be overwhelmed by SATOBs.

Table 2 also shows that sufficiently restricting the use of satellite data improved the 500-mb analysis significantly. Three subsets of data were established by combining RAOBs with allowed SATOBs based upon the 250-, 500- and 750-km restrictions. An analysis with the 250-km restriction reduced the first-guess 500-mb height errors by 34%, but the 500-mb temperature errors were slightly greater the first guess errors. The 500-km restriction was more effective; the analysis reduced the first guess 500-mb height errors by 48% and the 500-mb temperature errors by 45%. The 750-km restriction was a slight improvement; the analysis reduced the first guess 500-mb height errors by 50%, and the 500-mb temperature errors by 49%.

Table 2. The Root-Mean-Square Errors of a 500-mb Analysis of Geopotential Height (Z) and Temperature (T) in the Central American Region. The CONTROL analysis uses only rawinsonde observations, and the uncontrolled analysis uses both rawinsonde and satellite data. The restricted (Rstrctd) analyses are derived from observation sets from the satellite observations that were excluded if validated rawinsonde observations were available within the specified distances.

Element	First Guess	Control Analysis	Uncontrolled Analysis	Rstrctd 250 km	Rstrctd 500 km	Rstrctd 750 km
Z	35.7	17.4	38.1	23.7	18.5	17.8
T	2.16	1.06	2.16	2.16	1.19	1.09

The above results alone lead to the conclusion that the 750-km restriction appropriately prevents satellite data from degrading an analysis in rawinsonde data-poor regions; however, it is also possible to conclude that all satellite data should be eliminated. To gain further insight, the experiments were repeated so the results could be analyzed at individual rawinsonde stations. The specific purpose of the additional experiments was to determine how much satellite data were being incorporated into an analysis and its effect on the analysis at a point as a function of location and percentage of the number of total observations used to derive the analyzed value. The following discussion is limited to the analyses of geopotential height because of the similarity of results with the analyses of temperature.

In the uncontrolled (all SATOBs and RAOBs are available) analysis only 6 (of 67) analysis points had no SATOBs incorporated into the analysis; at 38 points the SATOBs were the only data used to perform the analysis. The analysis at the SATOBs-only points was better by an average of 4 m than the control (RAOBs only) analysis at only 8 points, while the analysis was worse by an average of 25 m at 28 points. These results show that relatively far away RAOBs are preferable to relatively close SATOBs. The largest errors (> 50 m) resulting from using SATOBs were distributed without preference either to land or to water.

At 20 points in the uncontrolled analysis both RAOBs and SATOBs were incorporated together for the analysis at those points. The combination of data types improved the analysis at 2 points by an average of 9 m and degraded the analysis at 15 points by an average of 32 m. Clearly, the effect on the analysis of including without restriction SATOBs with RAOBs was negative in this experiment. Consequently, a satisfactory specified-distance restriction had to be determined.

An analysis of geopotential height is arbitrarily defined at a point to be "significantly" improved if the error there compared to control is less than half the rms error of the controlled analysis of Z (which from Table 2 is approximately 9 m). Similarly, an analysis is significantly degraded if its error is increased by 9 m. Table 3 suggests that the 500-km restriction is better than either the 250- or 750-km restriction, both of which degrade the analysis much more than the 500-km restriction.

The effect of the restrictions is straightforward. The 250-km restriction was insufficient to stop SATOBs from unacceptably degrading the analysis. At 31 of the 67 analysis points, SATOBs were used in the analysis. The SATOBs had a significant effect at 19 of those 31 points, degrading the analysis at 15 points and improving the analysis at only 4 points. The 500-km restriction was much more effective. Nearly one-third of the points used SATOBs in the analysis, and 8 of those 22 points were significantly affected. The analysis was improved at 4 of the 8 points and degraded at the other 4. The 750-km restriction was ineffective. Only 11 points used SATOBs in the analysis, and 9 of them were significantly affected by SATOBs; however, the analysis was degraded at 8 of the 9 points.

Neither the points whose analysis was improved nor those whose analysis was degraded by incorporating SATOBs had a geographic preference for land, coastal, or ocean areas, although some of the largest analysis errors occurred well inland. The lack of geographic preference was unexpected. Even though additional experiments with different databases and methodologies for using SATOBs would be useful, time constraints caused acceptance of the above experiment's suggested 500-km restriction for use in the OSSEs.

Table 3. The Effect of Restricting the Use of SATOBs on an Analysis at 67 Rawinsonde Stations (whose observations had no influence on the analysis at that point). The only analysis points considered here are those that were significantly influenced by SATOBs.

Restriction To Using SATOBs	Points Using SATOBs	% of Points Improved	% of Points Unchanged	% of Points Degraded
none	20	0	50	50
250 km	19	21	0	79
500 km	22	18	64	18
750 km	11	18	9	73

2.2 STC TASK 2

Task 2 also had two subtasks. The primary subtask was to compute the first guess forecast errors, from which correlation functions could be developed. The secondary subtask was to determine and model observation errors, if these are needed by the OI scheme.

2.2.1 Modeling Observation Errors

The analysis from Section 2.1.1 showed that observation errors are not required in the OI procedure RAP uses. There was insufficient time during the OSSEs to confirm that accounting for observation errors was unnecessary in the RAP.

2.2.2 Forecast Error Correlations

The RAP interim report fully described the forecast error correlation databases. This subsection describes the structure of the (univariate and bivariate) forecast error correlations (horizontal, vertical, and temporal) and their modeling. The most obvious conclusion is that for all meteorological variables the univariate forecast error correlation structure is a strong inverse function of distance. In general, the forecast error correlations between a pair of observation points decrease exponentially with increasing distance between them out to about 1500 km, beyond which the data are essentially uncorrelated.

Recall that the correlations, which are the basis of the optimum interpolation scheme, are the correlation of the differences between forecasts and observations. The differences (or forecast errors) here are mostly a function of forecast length, season, and location. In general, the forecast errors (not necessarily their correlations, however) increase with forecast length, are larger in January than July, and are slightly larger in Eurasia than in central America (although the last is not clear cut because fewer observations are available for verification). Most of this analysis was performed using 36-hr forecast errors. The error correlations for the 36-hr forecast considerably exceed those for the 12-hr forecast and are much smoother out to a separation of 1000 km. Beyond 1000 km the 12-hr error correlations slightly exceed the 36-hr error correlations but probably not significantly.

2.2.2.1 The structure of spatial forecast error correlations at the 500-mb level

This section describes the distribution of average forecast error correlations of the meteorological elements (for example, A and B) organized according to relative direction and distance between

observations at points i and j (that is, $r_{ij}(\Delta A_i, \Delta B_j)$). Let temperature = T , dewpoint = T_d , humidity = Q , u- and v-component of wind = U and V , respectively, and height of mandatory geopotential level = Z .

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta Z_j)$ is slightly nonisotropic with a small preference (higher values) for the east-west orientation. The correlations were significant and positive out to 2000 km and dropped to zero at 2500 km.

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta T_j)$ is significantly positive out to only 700 km and drops to zero by about 1000 km. The structure is roughly isotropic.

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta(T_d)_j)$ is small but significantly positive out to 600 km to the north, 700 km to the east, 500 km to the south, and 400 km to the west. The correlation drops to zero in a pattern that suggests a nonisotropic structure; however, the lack of isotropy is probably not real. The best structure is to assume isotropy and let r decrease from 0.25 at 0 km to 0.1 at 700 km and 0.0 at 1000 km and beyond.

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta U_j)$ is notably different from the structure that resulted before the 45° axis translation and is physically meaningful. The change of orientation of the quadrants resulted in significantly positive correlations ($r \geq 0.25$) when point j is north of point i , and significantly negative correlations ($r \leq -0.30$) when point j is south of point i . Otherwise (that is, when the observation points lie along an east-west axis), the correlations are not significantly different from zero. To the north, however, the correlations are zero at nearby distances (< 200 km) and significantly positive from about 450 to 1650 km with a maximum ($r = 0.26$) at about 850 km. The structure to the south is nearly an inverse image, with a significantly negative core extending from 300 to about 1600 km that reaches a minimum ($r = -0.31$) at 800 km. In the north-south direction the correlation tapers to zero at about 2000 km. With these results and those from the preliminary analysis, the correlations can be assumed to be zero along the northeast-southwest direction and the northwest-southeast direction.

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta V_j)$ is consistent with the distribution of $r_{ij}(\Delta Z_i, \Delta U_j)$ (immediately above) except that the cores of significant correlations extend along the east-west direction, positive to the west and negative to the east. In the north-south direction r_{ij} is nearly zero.

The forecast error correlation $r_{ij}(\Delta T_i, \Delta Z_j)$ is similar to $r_{ij}(\Delta Z_i, \Delta T_j)$, which is described above.

The forecast error correlation $r_{ij}(\Delta T_i, \Delta T_j)$ is roughly isotropic and significantly positive out to about 1000 km and then drops to zero at 1200 to 1500 km.

The forecast error correlation $r_{ij}(\Delta T_i, (\Delta T_d)_j)$, also roughly isotropic, has small positive correlations for nearby points out to 400-500 km that drop to zero by about 800 km.

The forecast error correlation $r_{ij}(\Delta T_i, \Delta U_j)$ is consistent with $r_{ij}(\Delta Z_i, \Delta U_j)$, although its magnitude is not as large. The correlations are positive with j north of i , increasing to a maximum ($r = 0.21$) at 800 km and dropping to zero at 2000 km. When j is east of i or west of i , the correlations are zero at all distances. When j is south of i , the correlations are negative with a minimum ($r = -0.26$) at about 550 km that increases to zero at about 2000 km.

The forecast error correlation $r_{ij}(\Delta T_i, \Delta V_j)$ is zero at all distances when j is north or south of i . When j is east of i , the correlations are significantly negative from about 300 to 1400 km and then increase to zero at 2000 km. When j is west of i , the correlations are positive and significant from about 300 to 1400 km and then decrease to zero at 2000 km.

The forecast error correlation $r_{ij}[\Delta(T_d)_i, \Delta Z_j]$ is similar to $r_{ij}(\Delta Z_i, \Delta(T_d)_j)$, which is described above.

The forecast error correlation $r_{ij}[\Delta(T_d)_i, \Delta T_j]$ is similar to $r_{ij}(\Delta T_i, \Delta(T_d)_j)$, which is described above.

The forecast error correlation $r_{ij}[\Delta(T_d)_i, \Delta(T_d)_j]$ is significantly positive ($0.23 < r < 0.59$) out to about 600 km in the north-south direction, significantly positive ($0.12 < r < 0.49$) out to 800 km in the east-west direction, and tapers to zero beyond 1000 km.

The forecast error correlation $r_{ij}[\Delta(T_d)_i, \Delta U_j]$ is zero everywhere.

The forecast error correlation $r_{ij}[(\Delta T_d)_i, \Delta V_j]$, in contrast to $r_{ij}[\Delta(T_d)_i, \Delta U_j]$, is significantly negative when j is east of i from 500 to 1200 km, significantly positive when j is west of i from 350 to 500 km, and zero elsewhere. This pattern is consistent with $r_{ij}(\Delta T_i, \Delta V_j)$ and can probably be explained by the

climatological north-south gradient of T and T_a . Errors in v therefore have a much larger (and more significant) effect on errors of T_a than errors of u .

The forecast error correlation $r_{ij}(\Delta U_i, \Delta Z_j)$ is similar to $r_{ij}(\Delta Z_i, \Delta U_j)$, which is described above.

The forecast error correlation $r_{ij}(\Delta U_i, \Delta T_j)$ is similar to $r_{ij}(\Delta T_i, \Delta U_j)$, which is described above.

The forecast error correlation $r_{ij}(\Delta U_i, \Delta(T_a)_j)$ is zero everywhere.

The forecast error correlation $r_{ij}(\Delta U_i, \Delta U_j)$ is not isotropic. It is significantly positive in the north-south direction out to 750 km and tapers to zero at about 1000 km, and it is significantly positive in the east-west direction out to about 1000 km and tapers to zero at about 1600 km.

The forecast error correlation $r_{ij}(\Delta U_i, \Delta V_j)$ is zero everywhere.

The forecast error correlation $r_{ij}(\Delta V_i, \Delta Z_j)$ is similar to $r_{ij}(\Delta Z_i, \Delta V_j)$, which is described above.

The forecast error correlation $r_{ij}(\Delta V_i, \Delta T_j)$ is similar to $r_{ij}(\Delta T_i, \Delta V_j)$, which is described above.

The forecast error correlation $r_{ij}[\Delta V_i, \Delta(T_a)_j]$ is similar to $r_{ij}[\Delta(T_a)_i, \Delta V_j]$.

The forecast error correlation $r_{ij}(\Delta V_i, \Delta U_j)$ is zero everywhere.

The forecast error correlation $r_{ij}(\Delta V_i, \Delta V_j)$ is not isotropic. It is significantly positive in the north-south direction out to 1000 km and tapers to zero between 1300 and 1500 km; it is significantly positive in the east-west direction out to about 600 km and tapers to zero at about 1000 km.

The forecast error correlation $r_{ij}(\Delta Z_i, \Delta Z_j)_{\text{satellite}}$, $r_{ij}(\Delta Z_i, \Delta T_j)_{\text{satellite}}$, $r_{ij}(\Delta T_i, \Delta Z_j)_{\text{satellite}}$, and $r_{ij}(\Delta T_i, \Delta T_j)_{\text{satellite}}$ could not be calculated (see Section 2.2.3.2) by the methods used to calculate forecast error correlations from rawinsondes.

2.2.2.2 Forecast error correlations at 850- and 150-mb levels

The fine detail of the discussion at 500 mb will not be repeated here. Instead, only significant observations of the changes between the 500-, 850-, and 150-mb levels will be discussed. Regardless of the level, however, the general structure of the correlations is similar, that is, r usually decreases as the distance between observations increases. Nevertheless, there are some interesting differences.

For example, horizontal error correlations of height vs. height [$r_{ij}(\Delta Z_i, \Delta Z_j)$] at lower levels are normally larger than those at higher levels. This pattern, however, is not consistent with the other univariate correlations. The westerly wind error correlations in the east-west direction, which has the most significant correlations, have the opposite structure.

Figure 4 shows the average values of $r_{ij}(\Delta Z_i, \Delta Z_j)$ in the north-south direction (from quadrant 1 or 3) at 850, 500, and 150 mb. Only data that are significantly different from zero at the 95% level by Student's t-test are included in this and the figures that follow. Figure 5 shows the average $r_{ij}(\Delta U_i, \Delta U_j)$ as a function of distance between observation points i and j . Note that the correlations at the higher levels are greater than the correlations at the lower levels, the reverse of the pattern of $r_{ij}(\Delta Z_i, \Delta Z_j)$. These two examples show the importance of allowing the data to "speak for themselves" and resisting the temptation to simplify the structure.

Figure 6 shows clearly that the bivariate error correlations neither follow the univariate model's rapid decrease with distance nor have its distinct distribution in the vertical. For example, consider $r_{ij}(\Delta Z_i, \Delta U_j)$, whose characteristics at 500 mb were discussed above. Recall that the east-west error correlations were zero, whereas the north-south correlations were significantly positive to the north and significantly negative to the south. This pattern is also true at 850 and 150 mb; nevertheless, even though the differences between plots of bivariate error correlations at these levels are smaller than univariate correlations, the individual levels deserve the separate attention being paid to them.

2.2.2.3 Univariate and bivariate vertical error correlations

The vertical forecast error correlation of a meteorological element A on standard level k with meteorological element B on level l (k or $l = 1$ is 1000 mb, k or $l = 2$ is 850 mb, ..., k or $l = 12$ is 50 mb) at a rawinsonde observation point is expressed as $r_{kl}(\Delta A_k, \Delta B_l)$. For example, $r_{850,700}(\Delta A_{850}, \Delta B_{700})$, where A is on the 850-mb level and B is on the 700-mb level. The error correlations of Z-Z, T-T, U-U,

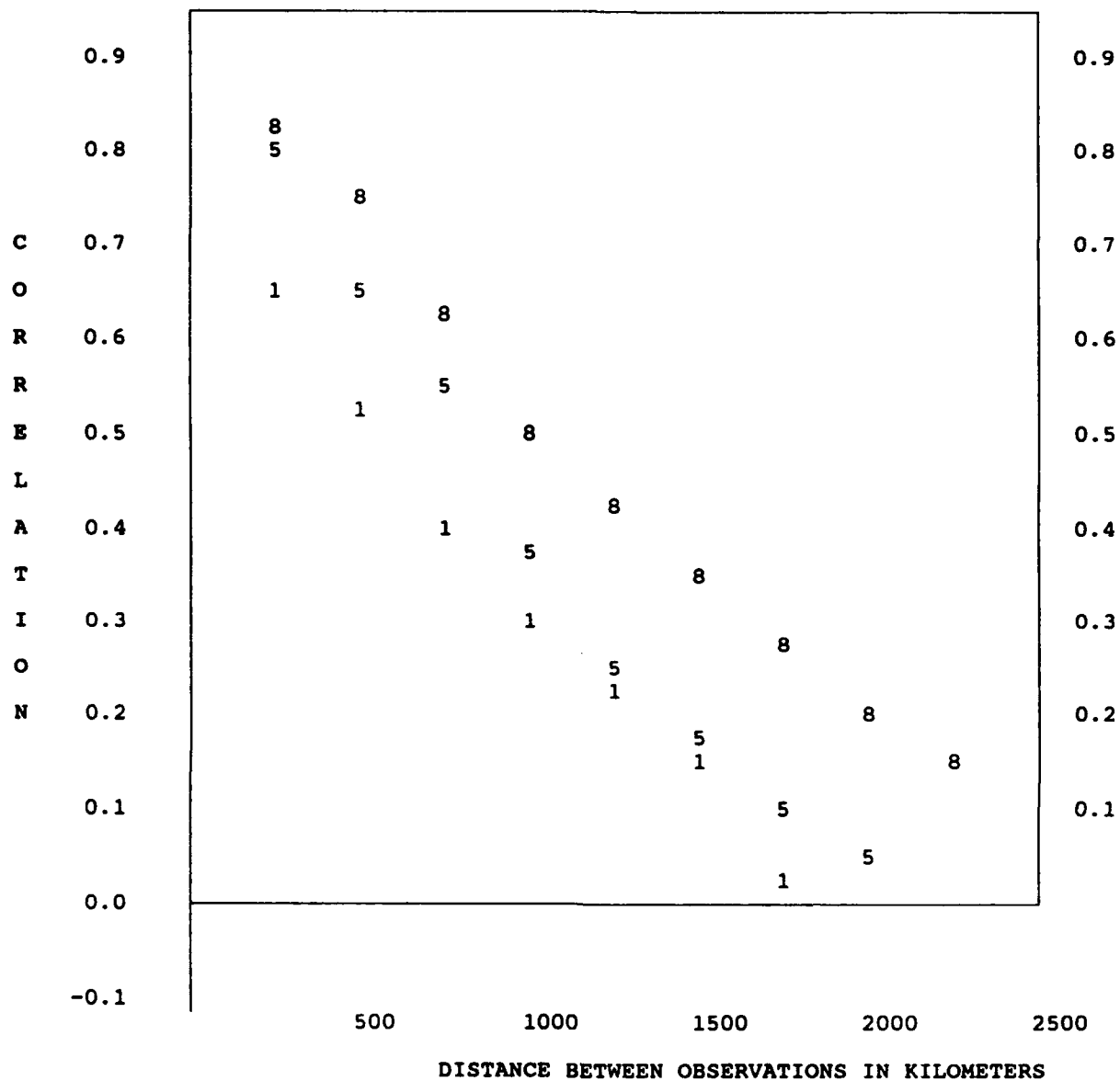


Figure 4. The $\Delta Z_i - \Delta Z_j$ error correlations along the north-south direction at 850 mb (8), 500 mb (5), and 150 mb (1), as a function of distance between observations.

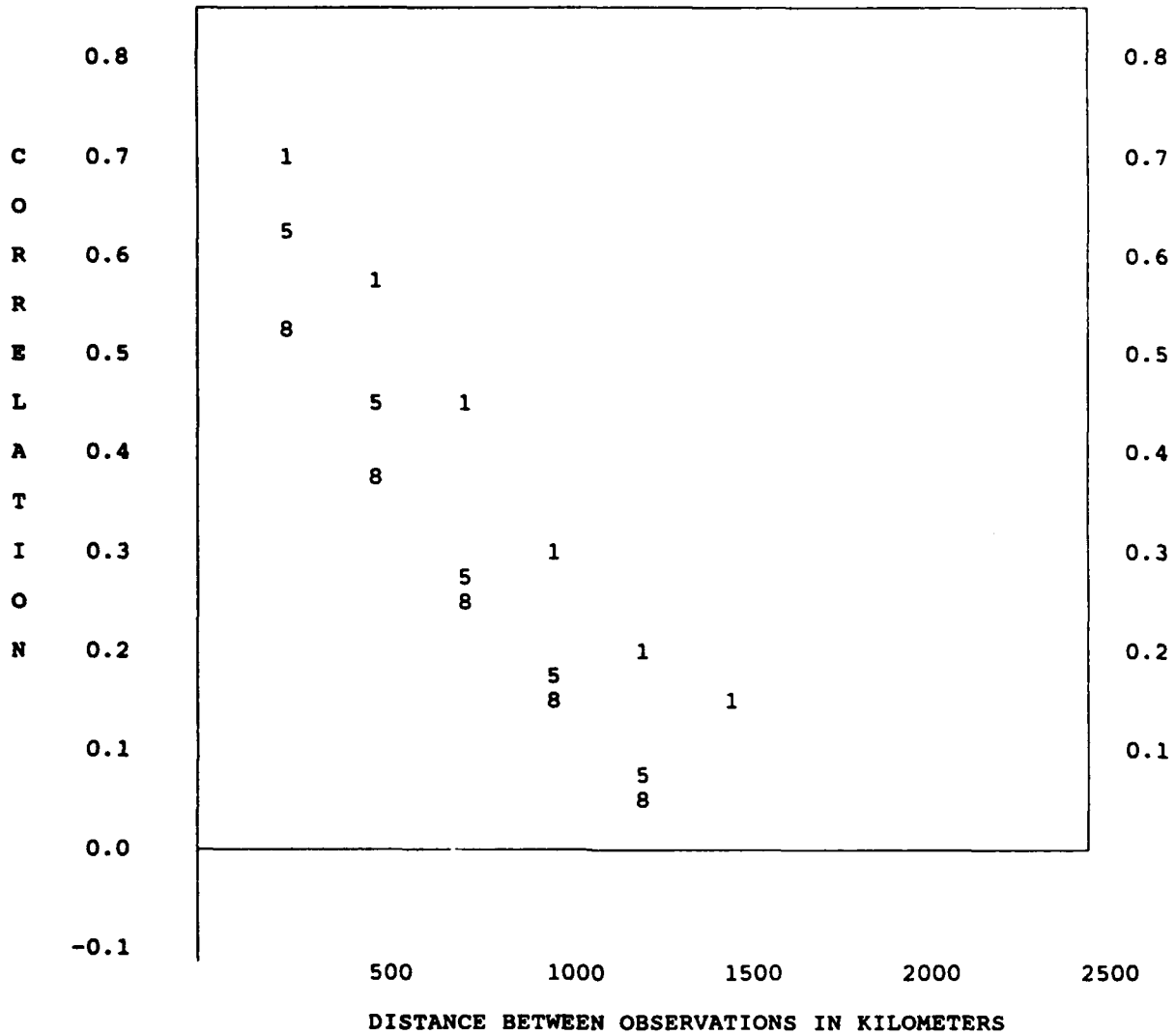


Figure 5. The average $\Delta U_i - \Delta U_j$ correlations $[r_{ij}(\Delta U_i, \Delta U_j)]$ from quadrant 2 (or 4) at 850 mb (8), 500 mb (5), and 150 mb (1), as a function of the distance between observations points i and j .

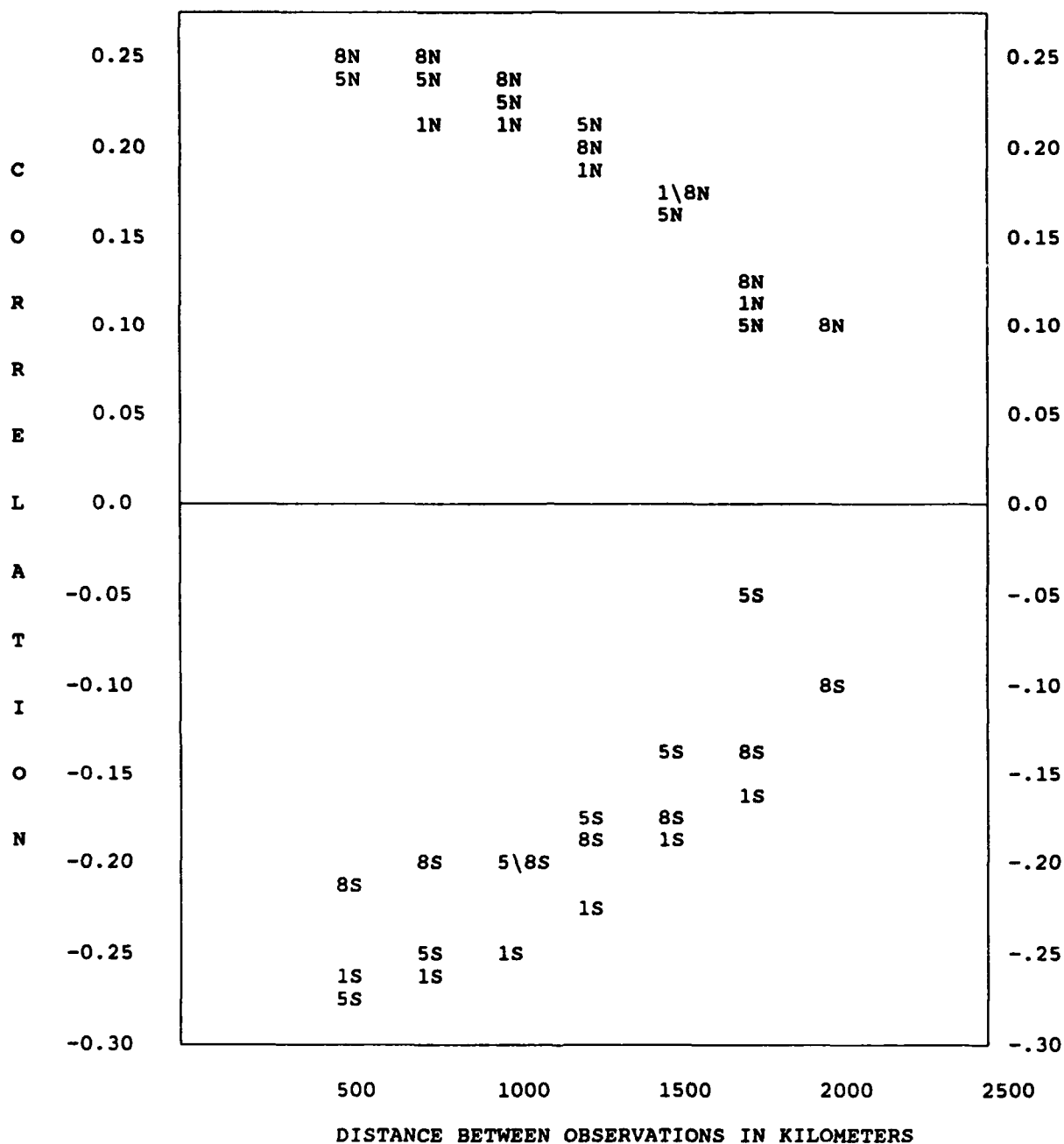


Figure 6. The average $\Delta Z_i - \Delta U_j$ correlations, $[r_{ij}(\Delta Z_i, \Delta U_j)]$, from quadrants 1 and 3 at 850 mb (8), 500 mb (5), and 150 mb (1), where the N or S following the 1, 5, or 8 identifies the direction (north or south) from observation point i to point j , as a function of the distance between observations.

V-V, Q-Q, Z-T, Z-U, Z-V, Z-Q, T-U, T-V, T-Q, U-V, U-Q, and V-Q were calculated and averaged for all rawinsonde sites in the regional databases for all mandatory levels at the given level with the levels above and below for each window (central America and Eurasia), month (January and July), and forecast hour (12, 18, 24, and 36). In addition, $r_k(\Delta A_k, \Delta B_l)$ is calculated.

The structure of the vertical error correlations in general is so complex that only complicated functions could describe them. Look-up tables can be developed using the results of the calculations and their significance according to Student's t-test. The following discussion is based on results from the Eurasian rawinsonde database of 36-hr forecast errors for January 1989. In practice, the value of $r_k(\Delta A_k, \Delta A_l)$, that is, the autocorrelations, from look-up tables would be used. The discussion below of the structure of the vertical error correlations refers to Table 4.

The forecast error correlation $r_k(\Delta Z_k, \Delta Z_l)$ is positive and statistically significant. The structure is simple at low and middle levels. The highest correlations at level k are at levels $l = k-1$ and $k+1$, where $r \geq 0.85$. The correlations gradually decrease as l moves further from k , as expected. The correlation of the lower and middle levels with the higher levels is small, apparently reflecting the "separation" of the troposphere from the stratosphere. This structure is one of the few vertical structures that could be represented by a relatively simple correlation function.

The forecast error correlation $r_k(\Delta Z_k, \Delta T_l)$ is clearly more complicated and so more typical. At $k = 1000$ mb, r is zero for $l = 1000$ to 300 mb; for $l = 250$ to 50 mb, r is (statistically) significantly negative but small, ranging between -0.086 and -0.162. At $k = 850$ mb, r decreases slowly from -0.115 at $l = 250$ mb to -0.187 at $l = 50$ mb; $r = 0$ at 300 mb; and r is about 0.1 below $l = 300$ mb. At $k = 700$ to 400 mb, r is similar to the case at $k = 850$ mb. At $k = 300$ mb, r is significant and nonzero for all l ; $0.163 \leq r \leq 0.594$ for l below 300 mb, and $-0.422 \leq r \leq -0.133$ for l above 300 mb. At $k = 250$ mb the structure is similar to $k = 300$ mb. At $k = 200$ mb, r is zero when $l = 250$ mb and when l is above 150 mb; r increases from 0.159 at $l = 1000$ to 0.507 at $l = 400$ mb, then decreases to zero at 250 mb, and is about 0.2 at $l = 200$ and 150 mb. At $k = 150$ mb the distribution is similar to $k = 200$ mb. At $k = 100$ mb the correlations are significantly positive for all l with only 1000 mb having $r \leq 0.20$. At levels $k = 70$ and 50 mb, r is zero when k is at the lower levels but significantly positive and increasing as k moves up the higher levels, reaching a maximum of 0.705.

Table 4. The Structure of the Vertical Correlations of Forecasts Errors Averaged over All Rawinsonde Sites at the Given Standard Level (k) with the Levels (l) Above and Below It for Each Window (Central America and Eurasia), Month (January and July), and Forecast Hour (12, 18, 24, and 36). The description of the data is provided below.

k	STANDARD LEVELS (mb)									
l										
1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
850	700	500	400	300	250	200	150	100	70	50
	850	850	850	850	850	850	850	850	850	850
	700	500	400	300	250	200	150	100	70	50
		700	700	700	700	700	700	700	700	700
		500	400	300	250	200	150	100	70	50
			500	500	500	500	500	500	500	500
			400	300	250	200	150	100	70	50
				400	400	400	400	400	400	400
				300	250	200	150	100	70	50
					300	300	300	300	300	300
					250	200	150	100	70	50
						250	250	250	250	250
						200	150	100	70	50
							200	200	200	200
							150	100	70	50
								150	150	150
								100	70	50
									100	100
									70	50
										70
										50

The forecast error correlation $r_{\mu}(\Delta Z_k, \Delta(T_d)_l)$ has weak correlations. At $k = 1000$ and 850 mb, r is zero for almost all l . From $k = 700$ mb up to $k = 300$ mb and l below 300 mb, r is significant but small ($0.093 \leq r \leq 0.234$).

The forecast error correlation $r_{\mu}(\Delta Z_k, \Delta U_l)$ is basically zero for all k and l .

The forecast error correlation $r_{\mu}(\Delta Z_k, \Delta V_l)$ has an odd distribution. At $k = 1000$, 850 , and 700 mb, r is significantly nonzero only when $l = k$ (r is about 0.10) or when $l \geq 100$ mb, where $-0.178 \leq r \leq -0.092$. At $k = 500$ up to 200 mb, r is significantly nonzero only when $l = 850$ to 400 mb, where $0.053 \leq r \leq 0.140$. For $k = 150$ mb and above, r is zero for all l .

The forecast error correlation $r_{\mu}(\Delta T_k, \Delta Z_l)$ is significant and nonzero for almost all levels. At $k = 1000$ to 700 mb, r is zero at only $l = 1000$ and 50 mb with maximum r ($0.14 \leq r \leq 0.55$) at $l = 1000$ to 700 mb and the five levels above these surfaces. At $k = 500$ and 400 mb, $r \geq 0.5$ for $l = k$ and the two levels above; r is near zero only at the two lowest and two highest levels. At $k = 300$ mb the pattern begins a transition. For all levels below ($l < k$), r is basically zero; at the levels above ($l \geq k$) $0.18 \leq r \leq 0.36$. At $k = 250$ and 200 mb, all r are highly significant, with negative r for $l \leq k$ and positive r above. The most negative r ($-0.42 \leq r \leq -0.33$) occurs with $k = 200$ mb and l at 250 down to 500 mb, that is, $r_{200,500-250}(\Delta T_{200}, \Delta Z_{500-250})$. At $k = 150$ mb and above, $r < 0$ for $l = 200$ mb and below. At $k = 100$ mb and above, r is strongly positive for $l \geq k$.

The forecast error correlation $r_{\mu}(\Delta T_k, \Delta T_l)$ is positive and well correlated for all $l = k \pm 1$. Up to $k = 300$ mb, $r \geq 0$ for $l \leq k$ and $r \leq 0$ for $l > k$ when l is below 100 mb. For $k = 250$ mb or above, $r \geq 0$ for $l \geq k-1$ and $r \leq 0$ for $l < k-1$. For $k = 100$ mb and above, $r \geq 0.58$ for $l = 100$ mb and above.

The forecast error correlation $r_{\mu}(\Delta T_k, \Delta(T_d)_l)$ has a simple structure (limited to 300 mb and below) with $r \leq 0.259$ for $k \neq l$. The correlations become smaller at higher levels and with increasing separation of levels.

The forecast error correlation $r_{\mu}(\Delta T_k, \Delta U_l)$ is zero for all levels above 1000 mb. At $k = 1$ (that is, at the 1000 mb level), the correlation with all levels l above, except 70 and 50 mb where r is zero,

is bounded by $0.15 < r < 0.36$. These correlations between ΔT_{1000} and $\Delta U_{850-100}$ are statistically significant and so are accepted as real.

The forecast error correlation $r_{kl}(\Delta T_k, \Delta V_l)$ has few correlations that are nonzero. The exceptions are at $k = 3-5$ (700-400 mb) where $r \approx 0.13$ for $l \geq 9$ (at 150 mb and above, and at $k = 10-12$ (100 mb and above) where $r \approx -0.14$ for $l = 4-7$ (from 500 up to 250 mb).

The forecast error correlation $r_{kl}[\Delta(T_d)_k, \Delta Z_l]$ has a complex distribution. At 1000 mb, ΔT_d is uncorrelated with ΔZ on the 1000-mb level but has a positive correlation ($0.10 \leq r \leq 0.17$) at all other levels. At 850 and 700 mb ($k = 2$ and 3), ΔT_d is negatively correlated with ΔZ at 1000 mb and positively correlated at 500 through 300 mb. At 500 and 400 mb ($k = 4-5$), ΔT_d is positively correlated with ΔZ only from the level below ($k = 1-1$) up to 300 mb ($l = 7$).

The forecast error correlation $r_{kl}[\Delta(T_d)_k, \Delta T_l]$ is quite different from $r_{kl}[\Delta T_l, \Delta(T_d)_k]$. The three lowest levels have a small correlation ($0.13 \leq r \leq 0.21$) for $k \neq l$, a slightly higher correlation for $k = l$, and a zero correlation for $l > k+2$. For 500, 400, and 300 mb ($k = 4, 5$, and 6), r is significant and positive ($0.12 \leq r \leq 0.24$) for $l = 2-5$, r is significant but negative ($-0.26 \leq r \leq -0.15$) for $l = 8$ and 9 , and r is zero elsewhere.

The forecast error correlation $r_{kl}[\Delta(T_d)_k, \Delta(T_d)_l]$ is significantly positive throughout most of the profile except the two lowest levels. At 1000 and 850 mb ($k = 1$ and 2), $r \approx 0$ for l above 500 mb and $r \approx 0$ for l below 700 mb. The correlations are particularly high for all $l = k \pm 1$, especially 300 mb. In general, r increases as k increases and l approaches k .

The forecast error correlation $r_{kl}[\Delta(T_d)_k, \Delta U_l]$ is zero above 700 mb for all l . The only significant correlations are at 1000 mb ($k = 1$) for levels $l = 1, 2, 3, 6$, and 7 , where $0.121 \leq r \leq 0.143$; and at 850 and 700 mb for levels $l = 1$ and 2 , where $0.091 \leq r \leq 0.176$. In general, r is too small to be of use to the OI procedure in practice.

The forecast error correlation $r_{kl}[\Delta(T_d)_k, \Delta V_l]$ is nonzero only for $k = 4$ and 5 , that is, the middle levels of 500 and 400 mb, where $0.100 \leq r \leq 0.161$ for $l = 2-6$.

The forecast error correlation $r_{kl}(\Delta U_k, \Delta Z_l)$ is zero for all k and l .

The forecast error correlation $r_{kl}(\Delta U_k, \Delta T_l)$ has an unusual distribution. For $k < 5$ (below 400 mb) the only significant correlations ($r \approx 0.20$) are at $l = 1$ (1000 mb). For $5 \leq k < 10$ the only significant correlations are $r \approx 0.20$ at $l = 1$ and $r \approx 0.11$ at $l = 9$. At the 100-mb level ($k = 10$), $r = 0.153$, -0.097 , and 0.114 for $l = 1, 6$, and 9 , respectively; r is zero for other l . At $k = 11$ (70 mb), $r = -0.128$ for $l = 7$, 0.11 for $l = 9$, and zero for all other l . At $k = 12$ (50 mb), $r = -0.14$ for $l = 7$ and 8 .

The forecast error correlation $r_{kl}(\Delta U_k, \Delta(T_d)_l)$ is weakly correlated at the three lowest levels and zero above. At levels $k = 1$ and 2 (1000 and 850 mb), $r \approx 0.15$ for $l = 1 - 3$ (from 1000 to 700 mb); r is zero for l above 700 mb. At $k = 3$, $r = 0.136$ at $l = 1$; $r \approx 0$ for all other l .

The forecast error correlation $r_{kl}(\Delta U_k, \Delta U_l)$ is well behaved, as are the other univariate correlations, with significantly positive correlations at almost all levels. The only exception is at 1000 mb where ΔU is uncorrelated with ΔU at 70 mb and 50 mb. The correlations increase with increasing k and are largest as l approaches k .

The forecast error correlation $r_{kl}(\Delta U_k, \Delta V_l)$ is zero most everywhere; the only significant correlations are at levels $k = 1 - 4$ and $k = 11$ and 12 , only for $l = 1$ ($0.145 \leq r_{kl}(\Delta U_k, \Delta V_l) \leq 0.317$). A strange exception is the $r_{12}(\Delta U_1, \Delta V_2) = -0.15$, the single significant negative correlation.

The forecast error correlation $r_{kl}(\Delta V_k, \Delta Z_l)$ also has a complicated distribution. The correlations are insignificant for $k = 1$, and 6 through 9 for all l . At $k = 2$, $r \approx 0.125$ for $l = 1 - 8$ and $r \approx 0$ for other l ; at $k = 3$, $r \approx 0.11$ for $l = 4 - 8$ and $r \approx 0$ for other l ; at $k = 3$, $r \approx 0.102$ for $l = 6$ and 7 , and $r \approx 0$ for other l ; and at $k = 4$, $r \approx 0.10$ for $l = 7$ and $r \approx 0$ for other l . At 70 and 50 mb ($k = 11$ and 12), r is not significant for $l > 3$ and $r \approx 0.15$ for $l = 1, 2$, and 3 .

The forecast error correlation $r_{kl}(\Delta V_k, \Delta T_l)$ is zero at $k = 1, 3, 7$, and 8 for all l . At $k = 2$, $r \approx 0.11$ for $l = 3$ and 4 , and $r \approx 0$ for all other l ; at $k = 4$, $r \approx 0.11$ for $l = 2$ and 3 , $r \approx -0.13$ for $l = 10 - 12$, and $r \approx 0$ for other l ; at $k = 6$, $r \approx -0.10$ for $l = 10$ and 11 , and $r \approx 0$ for all other l ;

at $k = 9$, $r \approx 0.13$ for $l = 1$, and $3 - 5$, and $r \approx 0$ for all other l ; at $k = 10$ and 11 , $r \approx 0.14$ for $l = 3 - 5$, and $r \approx 0$ for all other l ; and at $k = 12$, $r (= 0.162)$ is significant only at $l = 4$.

The forecast error correlation $r_k(\Delta V_k, \Delta T_k)$ is insignificant most everywhere except at 500 and 400 mb. The only nonzero values occur at levels $k = 2, 3, 4, \dots, 12$ only for levels $l = 4$ and 5 . These nonzero values increase from $r \approx 0.158$ at $k = 2$ to $r \approx 0.161$ at $k = 5$ and decreases to $r \approx 0.109$ at 50 mb.

The forecast error correlation $r_k(\Delta V_k, \Delta U_l)$ is zero for all k and l ($1 \leq k, l \leq 12$).

The forecast error correlation $r_k(\Delta V_k, \Delta V_l)$ has significant positive correlations, increasing as $l \rightarrow k$. The maximum r for $k \neq l$ occurs on the 300- and 250-mb levels where $r_{67} (= r_{76}) = 0.828$.

In general, all univariate vertical error correlations are significant and useful at nearly all levels. The bivariate vertical error correlations, while interesting from an analytical point of view, are small and most likely insignificant. Consequently, they are unlikely to be helpful to the OI scheme and so will not be used.

2.2.2.4 Temporal forecast error correlations

For meteorological elements A and B (recall that $A = B$ for univariate correlations), the temporal forecast error correlation is given by the general expression $r_{mn}(\Delta A_m, \Delta B_n)$, where m is the m th-hour forecast and n is the n th-hour forecast. The univariate and bivariate temporal forecast error correlation coefficients of all the meteorological elements (Z-Z, T-T, U-U, V-V, Q-Q, Z-T, Z-U, Z-V, Z-Q, T-U, T-V, T-Q, U-V, U-Q, and V-Q) were calculated and their averages stored. In addition, the multivariate correlations, say $r(A, B)$, were calculated also for both $r(A, B)$ and $r(B, A)$.

The four forecast times require six unique temporal forecast error correlations: $r_{\Delta t}: r_{12,18}(\Delta A_{12h}, \Delta B_{18h})$, $r_{12,24}(\Delta A_{12h}, \Delta B_{24h})$, $r_{12,36}(\Delta A_{12h}, \Delta B_{36h})$, $r_{18,24}(\Delta A_{18h}, \Delta B_{24h})$, $r_{18,36}(\Delta A_{18h}, \Delta B_{36h})$, and $r_{24,36}(\Delta A_{24h}, \Delta B_{36h})$. They are functions of the same variables as are the vertical correlations, that is, the rawinsondes, the months of January and July, the central American and Eurasian windows, and the mandatory levels.

Virtually all univariate temporal error correlations are significant and positive, decreasing as Δt becomes larger. The univariate temporal error correlations could be fitted to the function $e^{-C\Delta t}$, where C is a positive constant, determined from RWM forecast error statistics, and Δt is the absolute time difference either between two observations or between observations and analysis time. The bivariate error correlations cannot be generalized with such a simple model; look-up tables would be necessary to represent them.

A useful analysis can be made by grouping the six individual types of $r_{\Delta t}$ by level in terms of Δt as shown in Table 5, which is the model for the tables that follow. Note that Table 5 has two correlations at $\Delta t = 6$ and 12 hr because $r(\Delta A_{12h}, \Delta B_{18h})$ and $r(\Delta A_{18h}, \Delta B_{24h})$ are 6 hr apart and $r(\Delta A_{12h}, \Delta B_{24h})$ and $r(\Delta A_{24h}, \Delta B_{36h})$ are 12 hr apart. The tables with the data have the two 6-hr and 12-hr error correlations left-to-right rather than top-to-bottom.

Table 5. Model for Illustrating the Temporal Error Correlations Between Elements A and B on Standard Vertical Levels as a Function of $\Delta t = 6, 12, 18$, and 24 hr

level (mb)	$r_{\Delta t}(\Delta A, \Delta B)$ for Δt of			
	6 hr	12 hr	18 hr	24 hr
1000	$r(\Delta A_{12h}, \Delta B_{18h})$ $r(\Delta A_{18h}, \Delta B_{24h})$	$r(\Delta A_{12h}, \Delta B_{24h})$ $r(\Delta A_{24h}, \Delta B_{36h})$	$r(\Delta A_{18h}, \Delta B_{36h})$	$r(\Delta A_{12h}, \Delta B_{36h})$
850
700
.
.
70
50

The forecast error correlation $r_{\Delta t}(\Delta Z, \Delta Z)$ is well behaved and should be useful, especially for $\Delta t = 6$ hr, where $r > 0.53$ for all levels, and $\Delta t \leq 12$ hr, where $r > 0.35$ for all levels. At all levels r decreases approximately exponentially as Δt increases, most notably from 18 to 24 hr. The highest correlations are above 100 mb. Table 6 shows part of the data available for analysis.

The forecast error correlation $r_{\Delta t}(\Delta Z, \Delta T)$ is significantly positive above 150 mb, especially at 70 mb, and weaker but still significant on the 500-, 400-, and 300-mb levels. At the highest levels r increases slightly but erratically with increasing Δt . These bivariate correlations and $r_{\Delta t}(\Delta T, \Delta Z)$, which are larger, are the most significant of the bivariate correlations. Table 7 shows the actual correlations, where X means that the specific correlation was insignificant according to the Student t-test, probably because the sample size was too small. In practice, most of the values are too small to be of operational use to the OI procedure.

Table 6. Temporal Forecast Error Correlations ($\times 100$) of $r_{\Delta t}(\Delta Z, \Delta Z)$

Level	$\Delta t = 6$ hr	$\Delta t = 12$ hr	$\Delta t = 18$ hr	$\Delta t = 24$ hr
1000 mb	69 72	46 53	27	8
850 mb	59 60	44 59	30	12
700 mb	60 60	46 61	32	16
500 mb	59 60	39 57	29	14
400 mb	58 58	36 56	31	14
300 mb	56 48	37 55	31	17
250 mb	57 55	37 54	30	18
200 mb	53 56	39 53	34	22
150 mb	55 60	38 50	33	23
100 mb	56 60	42 52	39	33
70 mb	72 63	51 58	55	39
50 mb	72 69	57 63	63	53

Table 7. Temporal Forecast Error Correlations ($\times 100$) of $r_{\Delta t}(\Delta Z, \Delta T)$

Level	$\Delta t = 6$ hr	$\Delta t = 12$ hr	$\Delta t = 18$ hr	$\Delta t = 24$ hr
1000 mb	X	X	X	X
850 mb	X	10 X	18	9
700 mb	X	19 X	24	16
500 mb	19 X	18 X	23	14
400 mb	27 15	19 X	24	16
300 mb	X X	15 X	X	14
250 mb	X	X	X	X
200 mb	X	X	X	X
150 mb	X	X -9	X	X
100 mb	24 17	26 17	32	32
70 mb	56 55	41 39	54	44
50 mb	35 24	43 36	47	48

The forecast error correlations $r_{\Delta t}(\Delta Z, \Delta T_d)$, $r_{\Delta t}(\Delta Z, \Delta U)$, and $r_{\Delta t}(\Delta Z, \Delta V)$ are probably not useful because those few values that are significant are small ($r_{\max} < 0.20$). For purposes here these correlations are considered to be zero.

The forecast error correlation $r_{\Delta t}(\Delta T, \Delta Z)$ is more useful than $r_{\Delta t}(\Delta Z, \Delta T)$ because there are more significant correlations in general and they are higher. Table 8 shows that the most useful data are at $\Delta t \leq 12$ hr or at levels above 150 mb.

Table 8. Temporal Forecast Error Correlations (x100) of $r_{\Delta t}(\Delta T, \Delta Z)$

Level	$\Delta = 6$ hr	$\Delta = 12$ hr	$\Delta = 18$ hr	$\Delta = 24$ hr
1000 mb	X	X	X	X
850 mb	14 X	21 10	X	X
700 mb	19 X	26 14	X	X
500 mb	29 16	26 15	X	X
400 mb	26 23	31 17	X	X
300 mb	17 16	16 11	X	X
250 mb	X	X	X	X
200 mb	X	-18 -14	X	X
150 mb	X	X	X	X
100 mb	32 19	31 28	16	13
70 mb	42 43	48 44	30	33
50 mb	45 25	49 42	36	31

The forecast error correlations $r_{\Delta t}(\Delta T, \Delta T)$ is, like all univariate error correlations, well behaved. In general, the middle levels have the smallest correlations, the three highest levels have the largest correlations, and r decreases as Δt increases. An interesting exception occurs at the 1000-mb level (where the RWM forecast is usually poor), as shown in Table 9.

Table 9. Temporal Forecast Error Correlations ($\times 100$) of $r_{\Delta t}(\Delta T, \Delta T)$

Level	$\Delta = 6$ hr	$\Delta = 12$ hr	$\Delta = 18$ hr	$\Delta = 24$ hr
1000 mb	74 74	52 57	62	47
850 mb	48 45	29 39	15	15
700 mb	41 52	22 36	25	13
500 mb	37 48	22 36	20	11
400 mb	38 45	22 36	20	15
300 mb	37 35	25 21	X	12
250 mb	44 48	21 27	18	12
200 mb	55 59	34 38	19	X
150 mb	55 62	40 46	26	24
100 mb	56 66	48 60	49	42
70 mb	X 69	53 61	66	46
50 mb	X 55	47 58	54	43

The forecast error correlation $r_{\Delta t}(\Delta T, \Delta T_d)$ is significant only at 1000 mb, where the correlations decrease from 0.52 at $\Delta t = 6$ hr to 0.21 at $\Delta t = 24$ hr.

The forecast error correlations $r_{\Delta t}(\Delta T, \Delta U)$ and $r_{\Delta t}(\Delta T, \Delta V)$ are not significant considered to be zero for purposes here.

The forecast error correlations $r_{\Delta t}(\Delta T_d, \Delta Z)$ and $r_{\Delta t}(\Delta T_d, \Delta T)$ are insignificant.

The forecast error correlation $r_{\Delta t}(\Delta T_d, \Delta T_d)$ is significant at all levels, although the physical value of forecasts and measurements of moisture above the tropopause (approximately 300 mb) is open to question. Below 300 mb and above 850 mb, however, the correlations are significant only for $\Delta t \leq 12$ hr. On average $r \approx 0.31$ for $\Delta t = 6$ hr and $r \approx 0.15$ for $\Delta t = 12$ hr. At 850 and 1000 mb all r are significant. The highest values of r at 1000 mb where $r \approx 0.71$ at $\Delta t = 6$ hr and decreases to $r \approx 0.33$ at $\Delta t = 6$ hr. At 850 mb r is similar to levels above except that for $\Delta t > 12$ hr $r \approx 0.11$.

The forecast error correlations $r_{\Delta t}(\Delta T_d, \Delta U)$ and $r_{\Delta t}(\Delta T_d, \Delta V)$ are insignificant and taken to be zero.

The forecast error correlations $r_{\Delta t}(\Delta U, \Delta Z)$, $r_{\Delta t}(\Delta U, \Delta T)$, and $r_{\Delta t}(\Delta U, \Delta T_d)$ are insignificant and taken to be zero.

The forecast error correlation $r_{\Delta t}(\Delta U, \Delta U)$ is significant everywhere and has a uniform structure at all levels. The structure is similar to $r_{\Delta t}(\Delta Z, \Delta Z)$, shown in Table 6, although $r_{\Delta t}(\Delta U, \Delta U)$ appears to decrease almost linearly with increasing time rather than exponentially. In general, $r_{\Delta t}(\Delta U, \Delta U) < r_{\Delta t}(\Delta Z, \Delta Z)$ on all levels, as can be seen in Table 10.

The forecast error correlation $r_{\Delta t}(\Delta V, \Delta Z)$ is similar to $r_{\Delta t}(\Delta Z, \Delta V)$; some may be useful but probably not.

The forecast error correlations $r_{\Delta t}(\Delta V, \Delta T)$, $r_{\Delta t}(\Delta V, \Delta T_d)$, and $r_{\Delta t}(\Delta V, \Delta U)$ are insignificant and taken to be zero.

Table 10. Temporal Forecast Error Correlations ($\times 100$) of $r_{\Delta t}(\Delta U, \Delta U)$

Level	$\Delta = 6$ hr	$\Delta = 12$ hr	$\Delta = 18$ hr	$\Delta = 24$ hr
1000 mb	64 72	47 50	41	36
850 mb	44 50	31 37	25	17
700 mb	48 47	30 35	18	15
500 mb	51 55	35 41	27	17
400 mb	51 56	34 42	21	16
300 mb	53 57	35 45	24	14
250 mb	54 58	39 49	24	15
200 mb	51 58	37 53	29	18
150 mb	47 55	43 52	34	22
100 mb	54 62	49 53	36	37
70 mb	72 64	52 55	53	40
50 mb	64 64	53 61	58	47

The forecast error correlation $r_{\Delta t}(\Delta V, \Delta V)$ is similar to $r_{\Delta t}(\Delta U, \Delta U)$ in structure but slightly smaller for all Δt and levels.

Thus, similar to the univariate horizontal error correlations, the univariate temporal error correlations are typically statistically significant and have a structure that could be represented by analytic functions. The univariate temporal error correlations are highest for low and high levels at $\Delta t = 6$ hr. Without more data, the temporal bivariate error correlations in general would be useless in practice.

2.2.2.5 Summary of forecast error correlation models

The module to calculate horizontal forecast error correlations (Hollingsworth and Lonnberg [1986], Lonnberg and Hollingsworth [1986], and Thiebaut et al. [1986]) fits data to the models of two-dimensional (horizontal) bivariate forecast error correlations developed from rawinsonde data (Mitchell et al., 1990). The univariate vertical and temporal error correlations, however, are not modeled but represented by look-up tables containing statistically significant average correlations (or zero if insignificant). Forecast error correlations for satellite data will be shown to be similar to the rawinsonde data (Section 2.2.3), so the rawinsonde models will be used to represent satellite data as well.

The total error correlation of two elements can be treated as a separation of variables problem, and the four-dimensional correlation field assumed to be represented solely by space and time coordinates. That is, the total forecast error correlation, $r(x, y, p, t)$, can be modeled as the product of the three error correlations: $r(x, y) \cdot r(p) \cdot r(t) = r_{\Delta h} \cdot r_{\Delta p} \cdot r_{\Delta t}$ of all five meteorological elements on various levels, where $r_{\Delta h}$ ($= r_{ij}$) is the error correlation between observation points i and j on an isobaric level at a fixed time, $r_{\Delta p}$ is the vertical error correlation between standard vertical levels k and l (where $1 \leq k, l \leq 12$) at an observation point, and $r_{\Delta t}$ is the temporal error correlation between two forecasts (12, 18, 24, or 36 hr) at a point. When the total error correlation is calculated for use in the RAP OI procedure, only $r_{\Delta h}$ would be allowed to be either univariate or bivariate; $r_{\Delta p}$ and $r_{\Delta t}$ must be restricted to univariate. From a practical point of view, however, the OI procedure may be considered as virtually fully multivariate. In as much as the cross-variate correlations with respect to Δp and Δt are consistently near zero, any four-dimensional correlations involving such cross variables would not likely be selected by the stepwise regression procedure, except when no other observations are available.

2.2.3 Forecast Error Correlations of Satellite Data

A characteristic of satellite data (discussed below) resulted in the calculation of unrealistically high correlations compared with those obtained from rawinsonde data. As discussed in the RAP Interim Report, the rawinsonde method calculates the correlation between two observation points using a data sample consisting of 12 elements (12 forecasts 60-hr apart per month) per pair of observations. The correlations for all pairs are "binned" as a function of the distance and direction between the observations, and the average of all the correlations is calculated for each bin. These averages are then fitted to structure functions for models of the error correlations.

The satellite data required development of a separate error module. Since no fixed observation points exist in the satellite data set, the error data themselves—not the correlations—were binned as a function of the distance and direction between the observations. Then the correlations were calculated for each bin and fitted to the same structure functions as rawinsonde data. The error correlations of satellite data were similar to error correlations based on rawinsonde data only at 850 mb; at 700 mb and above the error correlations calculated by the satellite method were much higher.

The error correlations of rawinsonde data were recalculated, using the methodology of calculating correlations of the satellite data. The results of the two methods with identical rawinsonde data differed by a much greater degree than expected. Simulated correlation calculations with data sets of differing sizes determined the sensitivity of the two methods. The results with real data differed much more than the simulation tests.

Nevertheless, the rawinsonde method of calculating correlations between pairs of observation points, placing them into bins and calculating the average of the correlations in each bin, is the statistically appropriate methodology. The alternative (satellite) method, which combined all the forecast error into bins and then calculated the correlation, was found to be inappropriate due to the inherent nature of satellite data. A single pass of the satellite extends across the entire RAP window; thus a single pass of the satellite generally samples a wide range of values for any variable. Because of the method of binning the satellite data, each bin contains pairs of observations separated by a discrete interval. The observations, however, were not restricted to any particular location in the RAP window. Consequently, in any bin observations from southern and northern or eastern and western extremities of the region were intermingled.

This intermingling of data produced highly non-normally distributed data populations in the various bins because each bin contained error data from populations with nonhomogeneous characteristics. In particular, the variances would differ widely, especially for satellite passes with a long north-south extent. Chi-squared (X^2) tests demonstrated the lack of normality of satellite data (see Section 2.2.3.1). Simple examples can demonstrate that correlations calculated from non-normal distributions, such as the set of errors generated by the satellite method, would be much too high. Therefore, error correlations cannot be calculated for satellite data for the RAP region without increasing the size of the satellite data sample, since it would be necessary to restrict binning not only by observation, separation, distance, and direction, but also by geographic location. Hence, given the size of the data sample considered here, the error correlations developed from rawinsonde data will be used for all sensors (RAOBs, aircraft reports [AIREPS] and satellites).

2.2.3.1 Non-normal distribution of forecast errors

The failure of the "satellite" method (of grouping errors into bins as a function of separation distance and calculating a single correlation for that bin) to calculate forecast error correlations correctly is attributed to the non-normal distribution of errors over the Eurasian region. Numerical experiments based on the chi-square (X^2) test showed that the distribution of a set of errors assigned to a bin were unlikely to be normal.

Let an experiment have a variable distributed into h classes (intervals). Then define f_{oi} as the observed frequency of occurrence of the variable within the i th interval and f_{ei} as the theoretically expected frequency of occurrence within the interval, where

$$\sum_{i=1}^h f_{oi} = \sum_{i=1}^h f_{ei} = N \quad (24)$$

for N , the total number of observations.

A standard statistical theorem describes the X^2 test. For experiments with large N , $h \geq 5$, and $f_{ei} \geq 5$ for all i , the quantity X^2

$$\sum_{i=1}^h \frac{(f_{oi} - f_{ei})^2}{f_{ei}} \quad (25)$$

has approximately the X^2 distribution with $(h - k - 1)$ degrees of freedom, where k parameters from the theoretical distribution are computed from the experimental data.

Several numerical experiments were conducted for various pairs of meteorological elements, mandatory levels, and separation distances. Errors were consistently taken from each of the observation pairs used in the calculation of a correlation. To perform the test, the first step was dividing the range of errors into intervals (50 were used) and counting the frequency of occurrence of the errors within each interval. Intervals containing less than six occurrences were combined. Then two parameters of the normal distribution function (the mean and standard deviation) and the expected frequency of occurrence of errors within each interval were calculated, assuming a normal distribution. The number of degrees of freedom ranged from 29 to 38.

The hypothesis that the experimental data are normally distributed at the $\alpha\%$ level of significance was established. If

$$X^2 \leq X^2_{(h-k-1),\alpha} \quad (26)$$

the fit of the data to a normal distribution is good. Otherwise, the null hypothesis is accepted.

The inequality was not true (that is, the test of the hypothesis failed) for all the experiments at the 99.5% level of significance. The results indicated a non-normal distribution for height, temperature, and dewpoint. For winds the results were inconclusive. In no experiment, however, could the errors be shown to be from a normal distribution.

Figure 7 shows an example of how poorly the errors fit a normal distribution. In the graph it is clear that the tails of the data distribution appear to be normal but the middle is obviously not.

2.2.3.2 Justification for using the forecast error correlation model developed for RAOB data to model satellite data

The difficulty of calculating representative forecast error correlations with satellite data as a function of separation distance of pairs of observations needed to be resolved. The attempt to develop an independent correlation model for satellite data could not be cancelled unless the correlation models developed for RAOB data could reasonably be applied to satellite data. The contract manager suggested

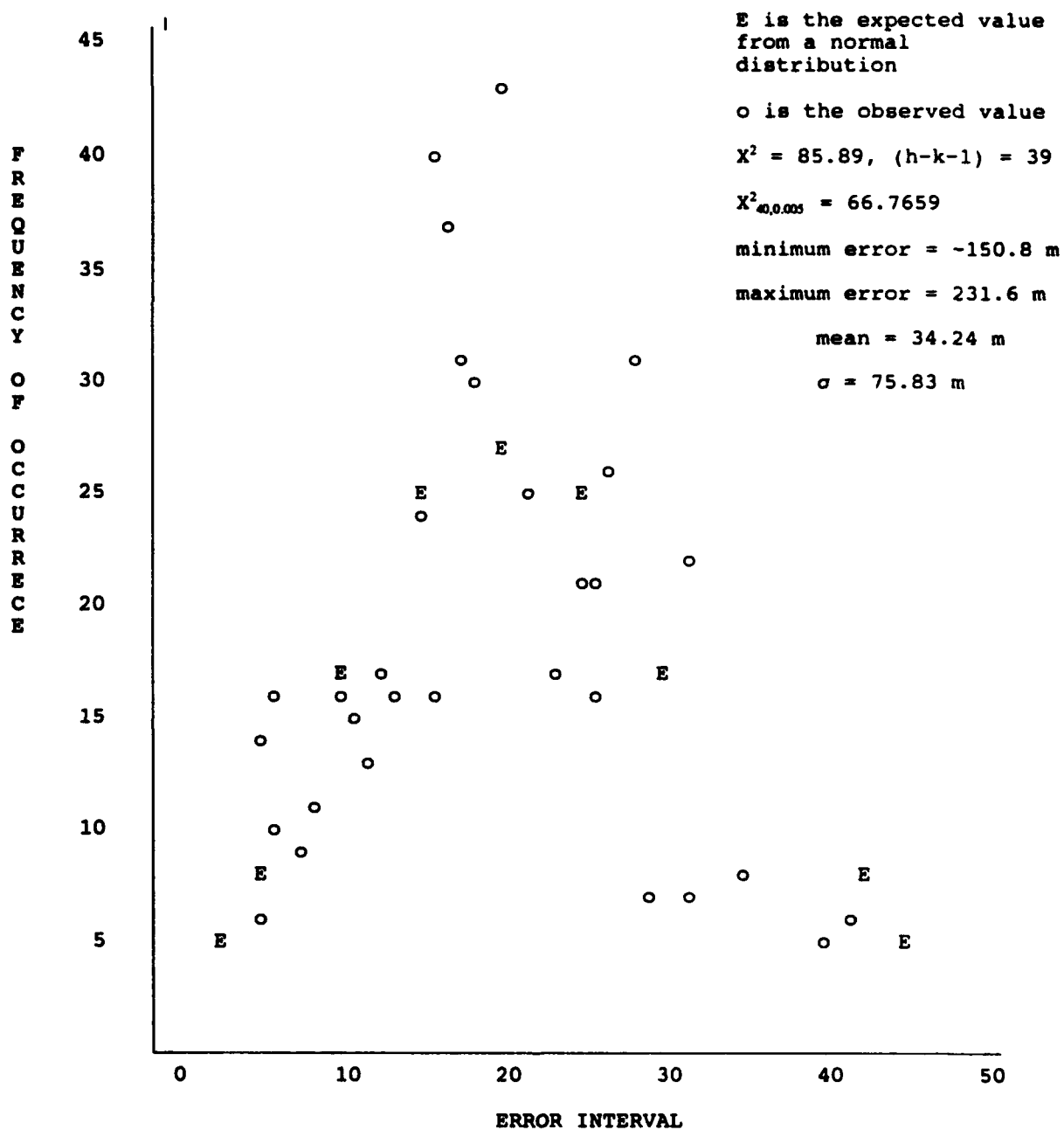


Figure 7. Approximate frequency distribution of 500-mb height errors from bin 10 (450 km < separation distance < 501 km).

another method of calculating error correlations of satellite data to show either that the assumption was valid or that independent error correlations should be calculated for the satellite data.

To use software that was already developed and thoroughly tested, STC performed numerical experiments with satellite data made to "look like" RAOB data. This involved interpolating satellite data to RAOB stations with the RAP OI procedure, and then calculating forecast error correlations as usual. Doing so allowed use of the established RAOB data correlation module and provided a baseline check with previously calculated error correlations. Furthermore, RAOB data were interpolated from surrounding stations to each RAOB station, and error correlations were calculated. These were compared with the error correlations developed from the actual data at each RAOB station.

Specifically, horizontal spatial forecast error correlations on three standard levels (850, 500, and 150 mb) were calculated between pairs of RAOB stations during January 1989 for three data types: observed RAOB data, interpolated RAOB data from surrounding stations to a station whose observation was withheld, and interpolated satellite data to those same RAOB stations. The error correlations for all station pairs were binned according to the separation distance of the pairs and averaged for each bin. A line to best fit the averages of each data type was drawn.

Figure 8 shows the approximate structure of the forecast error correlation of 500-mb height data as reported at RAOB stations (r), as interpolated to the RAOB stations (i), and as sensed by satellite and interpolated to RAOB stations (s). Clearly, the structures are highly correlated. Moreover, with $i > r$ for all separation distances, the interpolation procedure obviously increased the correlation.

To gain a quantitative understanding for the relationship between forecast error correlation models, consider the following analysis. Let R , R_i , and S be the average forecast error correlations calculated from pairs of observations binned according to their separation distance, where

- R was calculated from the observed data at RAOB stations as described in the interim report;
- R_i was calculated after interpolating RAOB data from surrounding stations to a given RAOB station, whose data were withheld from the analysis at that point;
- S was calculated from satellite data interpolated to RAOB stations;
- $r(S,R)$ is the correlation of forecast errors between satellite and RAOB data as a function of separation distance (100 to 2000 km) of observation pairs;

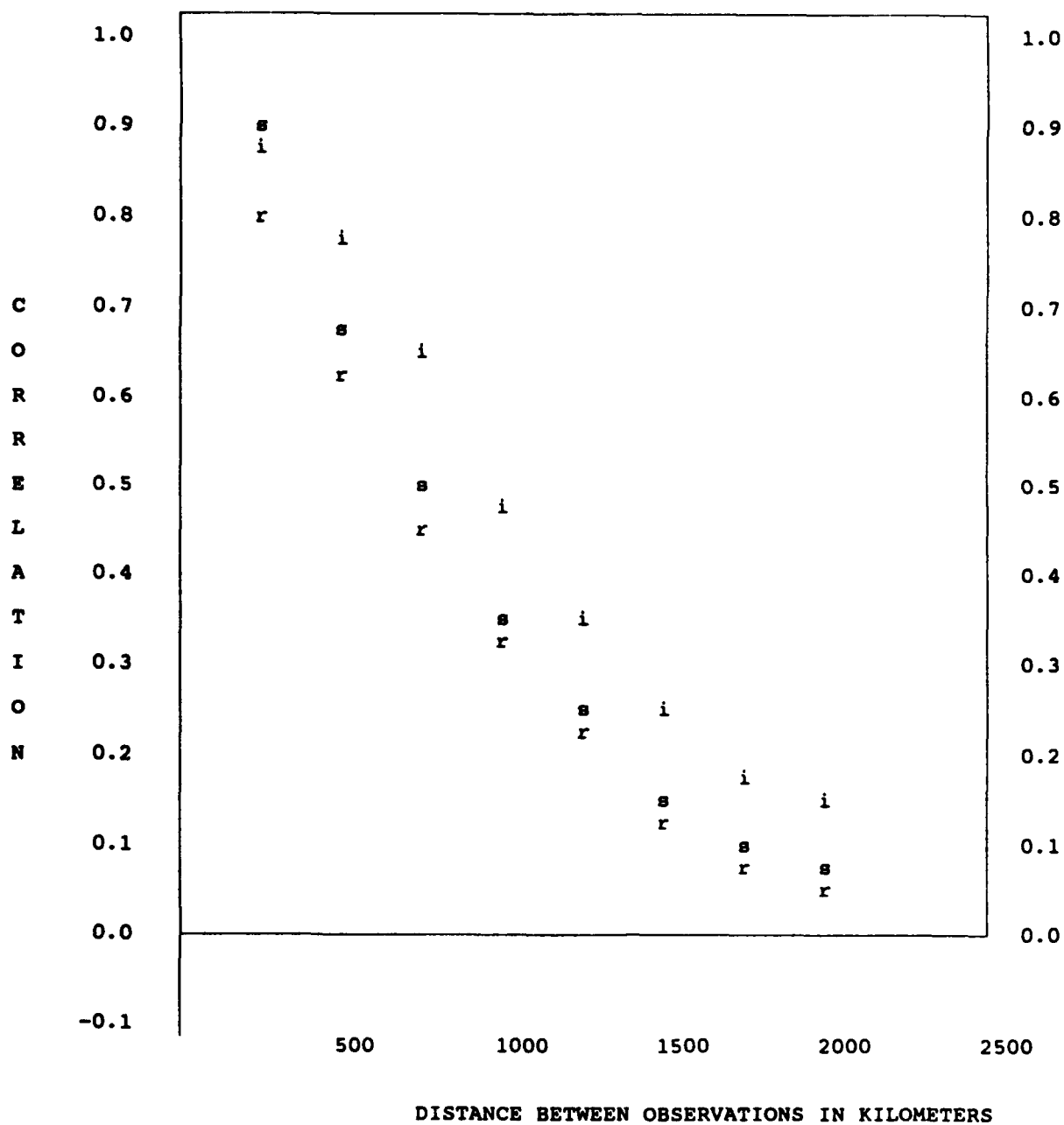


Figure 8. Height-height forecast error correlations at 500 mb in the north-south direction modeled from RAOB data (r), from interpolated RAOB data (i), and from interpolated satellite data (s).

\hat{S} is the linear regression estimate of the satellite error correlation function in terms of the observed RAOB error correlation function in the distance range from 100 to 2000 km;

\hat{R} is the linear regression of the estimate of interpolated R (i.e., R_i) in terms of the not-interpolated R ;

\bar{S} is the vertical average of S taken over 850, 500, and 150 mb; and

\bar{R} the vertical average of R taken over the same three levels.

With these definitions and data from Fig. 8, the following expressions for r , \hat{S} , and R can be readily calculated.

a. Height-height forecast error correlations

$$850 \text{ mb, } r(S, R) = 0.985 \text{ and } \hat{S} = -0.1015 + 1.15 \cdot R$$

$$500 \text{ mb, } r(S, R) = 0.997 \text{ and } \hat{S} = -0.0380 + 1.14 \cdot R$$

$$150 \text{ mb, } r(S, R) = 0.968 \text{ and } \hat{S} = -0.0214 + 1.21 \cdot R$$

$$\text{average: } r(\bar{S}, \bar{R}) = 0.972 \text{ and } \hat{S} = -0.0251 + 1.09 \cdot \Delta \bar{R}$$

$$500 \text{ mb: } r(R, R_i) = 0.974, \quad \hat{S} = -0.1043 + 0.998 \cdot R_i$$

$$r(R, R_i) = 0.986 \text{ and } \hat{R} = 0.083 + 1.10 \cdot R$$

b. Temperature-temperature forecast error correlations

$$500 \text{ mb: } r(S, R) = 0.983 \text{ and } \hat{S} = -0.090 + 1.20 \cdot R$$

$$150 \text{ mb: } r(S, R) = 0.998 \text{ and } \hat{S} = -0.129 + 1.00 \cdot R$$

c. U-wind forecast error components

$$500 \text{ mb: } r(S, R) = 0.999 \text{ and } \hat{S} = -0.044 + 1.19 \cdot R$$

Clearly, the structures of R and S are very closely related. In general, and especially with regard to temperature at 150 mb, the S correlations are slightly larger than the R correlations for the same distances. But as seen in Fig. 8 and in the statistical comparison of RAOB data (R) to interpolated RAOB data (R_i) in subparagraph a, the interpolated R is also higher than the observed R . The interpolation process apparently increased the correlation by about the same amount as the S correlation exceeded the R correlation.

Thus, there is a firm basis for using the error correlation functions developed for RAOBs also for satellite data. Nevertheless, other methodologies for calculating forecast error correlations with satellite data and larger data sets may reveal that improved correlation models could be developed for satellite data. No further work was performed toward this end, however.

2.2.4 The Development of Forecast Error Correlation Models

The general development of univariate forecast error correlation models and bivariate models involving wind follows from the work of Daley (1985), Hollingsworth and Lonnberg (1986), Lonnberg and Hollingsworth (1986), and Mitchell et al. (1990).

To derive the univariate and bivariate forecast error correlation structure functions involving wind components, it is convenient to use Helmholtz's theorem. That is,

$$\begin{aligned} u &= -\frac{\partial \Psi}{\partial y} + \frac{\partial X}{\partial x} \\ v &= \frac{\partial \Psi}{\partial x} + \frac{\partial X}{\partial y} \end{aligned} \quad (27)$$

where u and v are the zonal and meridional wind components, respectively, Ψ is the nondivergent stream function, and X is the irrotational velocity potential.

Let the forecast error correlations $r(\Delta_i, \Delta_j)$ be a function of the distance d between given meteorological variables at observation points i and j . Then the forecast error correlation equations can be written in polar coordinates as follows:

$$\begin{aligned} r(\Delta u_i, \Delta u_j) &= -\Gamma[F(d)] - \Phi[G(d)] + 2\Theta[H(d)] \\ r(\Delta v_i, \Delta v_j) &= -\Phi[F(d)] - \Gamma[G(d)] - 2\Theta[H(d)] \\ r(\Delta u_i, \Delta v_j) &= r(\Delta v_i, \Delta u_j) = \Theta[F(d)] - \Theta[G(d)] + \Lambda[H(d)] \\ r(\Delta a_i, \Delta u_j) &= -r(\Delta u_i, \Delta a_j) = \mathcal{Z}[I(d)] - \Pi[J(d)] \\ r(\Delta a_i, \Delta v_j) &= -r(\Delta v_i, \Delta a_j) = -\Pi[I(d)] - \mathcal{Z}[J(d)] \\ r(\Delta a_i, \Delta a_j) &= K(d) \end{aligned} \quad (28)$$

where a is a scalar variable (geopotential, temperature, or humidity),

$$\Gamma = R + \cos^2 \theta \, d^2 R^2, \quad \Phi = R + \sin^2 \theta \, d^2 R^2, \quad \Theta = \cos \theta \sin \theta \, d^2 R^2,$$

$$\Lambda = \cos^2\theta \sin^2\theta R^2, \Xi = \cos\theta dR, \Pi = \sin\theta dR,$$

and $\theta = (\theta_i + \theta_j)/2$ for θ_i and θ_j , the angles between the vector pointing due north and the direction vectors at observation points i and j . The direction vector at i is parallel to the great circle and points in the direction of least separation distance from point j ; the direction vector at point j is similar. The separation distance d_{ij} is the arc distance along the great circle between points i and j . Finally, R is the differential operator

$$R = \frac{1}{d} \frac{\partial}{\partial d} \quad (29)$$

that operates on the functions enclosed within [] in the system of Eq. 28.

The isotropic, homogeneous functions $F(d)$, $G(d)$, $H(d)$, $I(d)$, $J(d)$, and $K(d)$ from Eq. 28 are

$$\begin{aligned} F(d) &= \gamma_F r(\Delta\Psi_i, \Delta\Psi_j) \\ G(d) &= \gamma_G r(\Delta X_i, \Delta X_j) \\ H(d) &= \gamma_H r(\Delta\Psi_i, \Delta X_j) = \gamma_H r(\Delta X_i, \Delta\Psi_j) \\ I(d) &= \gamma_I r(\Delta a_i, \Delta\Psi_j) = \gamma_I r(\Delta\Psi_i, \Delta a_j) \\ J(d) &= \gamma_J r(\Delta a_i, \Delta X_j) = \gamma_J r(\Delta X_i, \Delta a_j) \\ K(d) &= r(\Delta a_i, \Delta a_j) \end{aligned} \quad (30)$$

where $\gamma_F = (E_\Psi/E_w)^2$, $\gamma_G = (E_X/E_w)^2$, $\gamma_H = E_X E_\Psi / (E_w)^2$, $\gamma_I = E_\Psi/E_w$, and $\gamma_J = E_X/E_w$. E_ξ is the rms error of the subscripted variable ξ , and w is a wind component such that E_w is either E_u or E_v .

The task of developing error correlation models is now reduced to determining the isotropic functions $F(d)$, $G(d)$, $H(d)$, $I(d)$, $J(d)$, and $K(d)$ from Eq. 30, which can be calculated from the forecast errors. First, however, to isolate the isotropic parts of the correlation models and to simplify the calculations, the wind components should be transformed from zonal, meridional (u, v) to transverse, and lateral (u_t, u_l) by

$$\begin{aligned} u_t &= -u \cos\theta + v \sin\theta \\ u_l &= u \sin\theta + v \cos\theta \end{aligned} \quad (31)$$

Applying Eq. 31 to the forecast error correlation models of Eq. 28 yields,

$$\begin{aligned}
 r(\Delta_i u_i, \Delta_j u_i) &= -RF - (d^2 R^2 + R)G \\
 r(\Delta_i u_i, \Delta_j u_i) &= -(d^2 R^2 + R)F - RG \\
 r(\Delta_i u_i, \Delta_j u_i) &= r(\Delta_i u_i, \Delta_j u_i) = -d^2 R^2 H \\
 r(\Delta_i u_i, \Delta_j a) &= r(\Delta_i a, \Delta_j u_i) = d RI \\
 r(\Delta_i u_i, \Delta_j a) &= r(\Delta_i a, \Delta_j u_i) = d RJ
 \end{aligned} \tag{32}$$

The forecast error correlations involving wind components, that is, the correlations on the left side of the system of Eq. 32, were calculated from the forecast error database using the methods described in Section 4.5.1 of the RAP Interim Report. An analysis of the results of experiments (using data from Eurasia in January 1989) shows that the following relationships are true: $r(\Delta_i u_i, \Delta_j u_i) \approx r(\Delta_i u_i, \Delta_j u_i) \approx r(\Delta_i u_i, \Delta_j a) \approx r(\Delta_i a, \Delta_j u_i) \approx 0$, which in turn requires H and J to be approximately zero. Moreover, G has been shown to be much less than F, which alone has provided a reasonable fit of the data; so the correlation models involving wind components will be nondivergent ($G \approx 0$).

Thus, the only remaining (nonzero) functions from Eq. 32 are F and I, which are the forecast error correlations of wind components

$$r(\Delta_i u_i, \Delta_j u_i) = -RF \text{ and } r(\Delta_i u_i, \Delta_j u_i) = -(d^2 R^2 + R)F, \text{ where}$$

$$F(d) = \frac{\beta}{1 + \alpha} [f(c, d) + \alpha f(\frac{c}{N}, d)] \tag{33}$$

and

$$f(c, d) = (1 + cd + \frac{c^2 d^2}{3})e^{-cd} \tag{34}$$

from Mitchell et al. (1990). The parameters $\alpha = 0.2$ and $N = 3$ are chosen to improve the fit of the data to the curves, while the empirical constants β and c are determined by using the IMSL subprogram RNLIN to provide an equation of best fit of the forecast error correlations to the model.

The other remaining functions, I from Eqs. 32 and K from Eq. 30, are treated in a similar fashion to provide the other forecast error correlation models, which also require calculation of empirical constants β' and c' .

2.2.5 Fitting Real Data to the Forecast Error Correlation Models in Practice

The application of the above theory occasionally yielded poor results in practice. Whenever the distribution of the correlations was widely scattered or sharply varying, RNLIN failed to develop a proper curve of best fit. For example, all the models of 18-hr forecast error correlations failed, probably due to insufficient data available from offtime observations. The philosophy of "letting the data speak for themselves" was appealing but difficult to implement in automated production. In addition, modeling all forecast error correlations was especially difficult in the stratosphere, probably because less data are available there than in the troposphere. Consequently, additional work beyond the scope of this project was necessary to have developed better models.

The possible maximum of 5,200 forecast error correlation models would have overwhelmed development efforts and resulted in many trivial (null) models anyway. The small sample of data was inadequate for the task of assuring statistically significant correlations; consequently, all insignificant correlations have been modeled as zero. The effect was to truncate most models at approximately 1000 and almost all at 1500 km. Moreover, except for the models of geopotential and wind components, the bivariate models in the stratosphere are zero for all observation separation distances.

The above approach was a result of analyzing the structure of hundreds of forecast error correlations as a function of observation separation distance (Section 2.2.2.1). The models, as represented by curves of best fit, therefore also account for sound physical principles of weather analysis. Portions of some curves, even though passing the statistical significance test, were set to zero because their structures made no physical sense. Usually this "zeroing" of a model occurred when the curve, decreased as an exponential function of distance (which is typical) to zero at approximately 1000 km or so, became negative, and then turned positive and significant beyond 1500 km or more. Correlation curves such as this are rejected for two reasons. Observations 1500 km or more beyond a point will not be used to provide an analysis there; therefore, correlations outside that range should not be part of the model. Also, no matter what reasonable level of statistical significance is chosen, some random distribution of data will have a "significant" correlation.

The optimum interpolation analysis technique relies on the capability to develop physically realistic error correlation models, which in effect determine those observations that will influence an analysis at a point and how strong that influence will be. The horizontal forecast error correlation

coefficients were fitted to structure functions as described in Section 2.2.4, and were modeled from the average correlation calculated as a function of the distance between observation pairs.

Several experiments demonstrated the sensitive nature of the forecast error correlation models, as expressed by the system of r_{ij} in Eq. 6 when solving for the weights, to inaccurate intercorrelations. Often the system of equations was on the edge of numerical instability because the intercorrelations (usually those involving u- and v-wind components) were dynamically inconsistent, presumably due to either inaccurate forecast wind error models or an inadequate sample size of data used to fit the model, or both.

These experiments resulted in a poor analysis at any point where the FSR observation selection procedure failed to stop accepting more observations, eventually including those making no physical sense, because the multiple correlation continued to increase (in extreme cases exceeding unity). The weights, calculated from Eq. 12, associated with these large sets of observations provided an analysis with unacceptable errors. A diagnostic study of the evolving observation selection process revealed that the first three or four selections would have provided a good analysis (see Section 2.1.1).

Thus, the problem could have been resolved from at least two possible approaches, either changing the forecast error correlation models or modifying how they were employed in the OI scheme, or a combination of the two. (A much larger dataset for producing the forecast error correlation models, however, probably would have provided more representative, consistent intercorrelations resulting in better models, thereby eliminating the problem.)

One pragmatic approach was simplifying the wind error correlation models by applying the systematic analysis described in Section 2.2.2.2. Basically, this involved accepting only those correlations that passed Student's t-test and modeling the correlations as zero elsewhere. There was little effect on the univariate models; however, many bivariate models became null. The simpler models resulted in fewer failures of an analysis at a point.

Finally, the concept of modeling the error correlations by quadrants (to account for anisotropy) probably caused the OI scheme to fail. The distribution of observations, typically into all four quadrants about a point, brought four different correlation models into play. When the direction vector connecting

observations crossed from one quadrant into another, the resulting correlation matrix had inconsistent correlations, which caused analysis failures. Averaging the data from all quadrants eliminated some of the inconsistent intercorrelations; however, averaging also turned the four correlation models into a single, isotropic model. But the forecast error correlations are anisotropic, especially those among geopotential and wind components and the wind components themselves (see Section 2.2.2.2).

The partial solution to the problem was to return to a method, abandoned earlier when its need was unclear, of representing the wind velocity vector by its transverse and longitudinal components. Buell (1972) took the earth-coordinate-oriented wind components (u, v) and transformed them to transverse and longitudinal components (u_t and u_l , respectively) by a simple rotation of coordinates (Eq. 31). The longitudinal component (u_l) is the component in the direction from observation point 1 to observation point 2, and the transverse component (u_t) is at right angles to the direction (positive to the left when looking) from observation point 1 to observation point 2.

The RAP forecast error correlation models therefore used the system of equations Hollingsworth and Lonnberg (1986) developed for the statistical structure of the forecast error correlations in the wind field, following Buell (1972) and Daley (1985), and using u_t and u_l . These expressions (Eq. 32) are internally self-consistent, relate the forecast error correlations of geopotential with the forecast error correlations of the velocity, and allow the isotropic component to be isolated and modeled. The resulting error correlation models fit the data better than models based upon u and v and provided intercorrelation matrices that were more consistent.

Nevertheless, in the practice of the OI procedure, the stability of the multivariate matrixes from Eq. 8 was assured only by the use of Eq. 14. This constraint imposed by Eq. 14 is likely due to imperfections in the models that resulted in inconsistent intercorrelations.

2.3 STC TASK 3

The purpose of Task 3 was to document formally the technical quality of RAP. This required the development of a relocatable verification package, using standard measures of error such as the RWM verification model (RWMVER), the Phillips Laboratory (PL) Global Spectral Model (GSM) diagnostic package, and map comparisons. The verification package tested RAP in real and simulated data experiments, that is, comparing RAP to the AFGWC High-Resolution Analysis System (HIRAS) model

and real observations, and with "truth" in the observing systems simulation experiments. Note that the verification experiments were not designed to simulate an operational test; they were designed to "prove the RAP concept."

The first phase of STC Task 3, modifying the AFGWC Relocatable Window Forecast Model verification model for the RAP system operations, was completed in late 1991. After additional modifications that made the RWMVER better suited for RAP, the package was used to test RAP formally in real data experiments. The results are discussed in Section 2.3.2.

2.3.1 Verification of RAP Analyses

The RWMVER was not used, however, as the verification method to determine either the capability of RAP to perform analyses² or the quality of RAP's forecast error correlation models. As discussed in the RAP Initial Work Plan (1989), the OI procedure is performed at observation points when assessing the results of numerical analysis experiments. The observed values at a given point are withheld from the analysis when the analysis is being performed there, and the analyzed values are then compared to the observed values. This method of verification is better than simply interpolating the values from the four surrounding grid points to an observation point and comparing the interpolated value to the observed value, whose large weight strongly affected the analysis at the surrounding grid points.

Experiments using the forecast error correlation models developed with regional data and multivariate optimum interpolation techniques yielded results that were superior to early experiments that used a simple correlation model with a univariate OI scheme (fully described in Section 5 of the RAP Interim Report). The 36-hr forecast error correlation model was chosen to illustrate RAP's effectiveness even when using the least accurate first guess forecast.

The multivariate analysis is similar to the univariate analysis; however, the forecast error correlation models from STC Task 2 yielded an improvement for geopotential, temperature, and zonal

² The RWMVER simply interpolates values from the four surrounding grid points to an observation point and compares the interpolated value to the observed value, whose large weight strongly affected the analysis at the surrounding grid points. In effect, RWMVER (whose primary purpose is to verify RWM forecasts, not compare an analysis to observations) essentially compares an observation to itself when performing grid-to-observation verification.

wind. On the other hand, the models of the dewpoint and v-wind component were similar to the univariate model described in the interim report, as seen in Tables 11a, 11b, and 11c. These tables show the rms differences of analyses performed at 850 mb, 500 mb, and 300 mb over Eurasia on 12 January 1989 at 1200 UTC. This data set was carefully checked for observation errors both manually and with the buddy check.

This experiment, whose results were confirmed by other experiments, points out the need for improved forecast error correlation models, which could be developed from more data. At 850 mb the results of this experiment were not as good as those at either 500 mb or 300 mb. Nevertheless, all analyses show positive skill (that is, analysis errors are smaller than first guess forecast errors). One of the most interesting results is that the univariate models are nearly as effective as the bivariate models, suggesting that the OI procedure could be made much less complicated and still retain its skill at performing analyses. An alternative conclusion, however, is that better bivariate models should be developed to improve analyses further.

Table 11a shows that the multivariate scheme had only a small effect. Very few multivariate observations were selected, probably because the cross correlations were so low. The forecast error correlation models are only a slight improvement over the simple model used earlier except for the model of geopotential; the dewpoint model is not as good. But the relatively poor performance of the dewpoint model may be related to the large first guess forecast error in this case.

From Table 11b, however, it is clear that the analysis scheme is an unqualified success for geopotential, temperature, and the u-wind component, even though the analyzed dewpoint and the v-wind component were similar to those provided by a simple model. This is perhaps to be expected with dewpoint, which is a difficult variable to analyze and whose model is less reliable than the others, but not with the v-wind component. In this case neither the univariate nor the bivariate model of the v-wind forecast error correlations paid dividends for the effort to develop them.

Table 11a. Root-Mean-Square Differences from Withheld Radiosonde Values at 850 mb of the First Guess Forecast, an Univariate Analysis with a Simple Correlation Model, a Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis

850 mb

Analyzed Variable	First Guess Forecast (36 hr)	Simple Univariate Models	Data-Fitted Univariate Models	Data-Fitted Bivariate Models
Geopotential (m)	63.0	27.6	18.4	19.7
Temperature (K)	3.29	2.43	2.23	2.35
Dewpoint (K)	6.16	4.79	4.97	4.97
U-wind (m/s)	5.18	4.05	3.88	3.86
V-wind (m/s)	5.02	3.56	3.37	3.37

Table 11b. Root-Mean-Square Differences from Withheld Radiosonde Values at 500 mb of the First-Guess Forecast, an Univariate Analysis with a Simple Correlation Model, an Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis

500 mb

Analyzed Variable	First Guess Forecast (36 hr)	Simple Univariate Models	Data-Fitted Univariate Models	Data-Fitted Bivariate Models
Geopotential (m)	73.0	35.8	21.0	20.9
Temperature (K)	2.88	1.94	1.42	1.36
Dewpoint (K)	7.65	5.63	5.73	5.73
U-wind (m/s)	7.48	5.20	4.26	4.25
V-wind (m/s)	6.62	4.49	4.44	4.44

The results at 300 mb (Table 11c) are similar to those at 500 mb; geopotential and wind are analyzed very well but not dewpoint temperature. All models developed from data, however, are superior to the simple univariate model (an exponential function). Nevertheless, the bivariate models appear to add little to the information content of the univariate models.

In summary, these experiments used the forecast error correlation models developed with regional data and employed multivariate optimum interpolation techniques, including observation selection by a stepwise regression procedure. The results were superior to prior experiments that used a simple correlation model with a univariate OI scheme (fully described in Section 5 of the Interim Report). The multivariate and univariate analyses were similar, although the forecast error correlation models from STC Task 2 yielded an improvement for all analyzed variables.

Although only one case study is considered here, additional experiments yielded similar results. The overall verification program would have been more rigorous if the results of several experiments, using independent observation sets, could have been combined. There was insufficient time to perform additional experiments, however, and independent data sets were unavailable. In any event, the results obtained satisfied the purpose of the task, that is, to verify the goodness of RAP and point out where more research and development are needed.

Table 11c. Root-Mean-Square Differences from Withheld Radiosonde Values at 300 mb of the First-Guess Forecast, an Univariate Analysis with a Simple Correlation Model, an Univariate Analysis Derived from Correlation Models Developed from Real Data, and a Multivariate Analysis

300 mb

Analyzed Variable	First Guess Forecast (36 hr)	Simple Univariate Models	Data-Fitted Univariate Models	Data-Fitted Bivariate Models
Geopotential (m)	88.9	54.4	46.7	45.0
Temperature (K)	2.34	2.10	2.14	2.14
Dewpoint (K)	5.96	4.64	4.91	4.92
U-wind (m/s)	10.9	7.91	6.65	6.53
V-wind (m/s)	9.48	6.79	6.62	6.62

2.3.2 Comparison of RAP Analyses with HIRAS Analyses

This section describes the results of comparing analyses derived from RAP with analyses derived from HIRAS. The results are from a uniform gridded data field (UGDF) grid-to-observation verification, which has RWMVER's technical weaknesses described in Section 2.3.1 (footnote 2). For this and other reasons these results must be interpreted carefully. The RAP analyses were performed on a higher resolution grid than the HIRAS grid, giving RAP an advantage. Finally, the sample sizes are small. Nevertheless, the tables shown below illustrate RAP's unequivocal superiority over HIRAS. The results of the verification experiments above combined with the comparisons here prove the validity of the "RAP concept."

The tables show comparison of the rms errors at 850, 500, and 300 mb of RAP analyses to the rms errors of HIRAS. The analysis variables are geopotential (Z), temperature (T), dewpoint temperature (T_d), whose analysis errors are more easily interpreted than relative humidity errors, the zonal wind component (U), and the meridional wind component (V). Tables 12a, 12b, and 12c show RAP analyses were derived from 12-hr forecast error correlation models and Tables 13a, 13b, and 13c show similar analyses derived from 36-hr forecast error correlation models. The Samples column shows the number of observations used in the verification; The Analysis Region is either Eurasia (EUR) or central America (CAM); the variables T, Z, RH, U, and V are subscripted by either R if analyzed by RAP or by H if analyzed by HIRAS. The results are shown as the ratio of RAP analysis errors to HIRAS analysis errors. Note that RAP errors are always much less than HIRAS errors, sometimes by an order of magnitude.

Table 12a. The Ratio of the Root-Mean-Square Errors of 850-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 850-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
1 Jul 88	99	EUR	0.25	0.10	0.09	0.09	0.09
1 Jan 89	86	EUR	0.25	0.22	0.11	0.06	0.15
4 Jan 89	100	EUR	0.17	0.13	0.10	0.18	0.12
6 Jan 89	100	EUR	0.28	0.27	0.12	0.13	0.10
1 Jan 89	23	CAM	0.05	0.09	0.18	0.02	0.06
4 Jan 89	21	CAM	0.21	0.05	0.09	0.03	0.03
6 Jan 89	32	CAM	0.08	0.07	0.09	0.02	0.07

Table 12b. The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 500-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
1 Jul 88	88	EUR	0.11	0.22	0.18	0.10	0.09
1 Jan 89	81	EUR	0.21	0.11	0.14	0.07	0.08
4 Jan 89	100	EUR	0.18	0.10	0.23	0.07	0.10
6 Jan 89	96	EUR	0.30	0.34	0.12	0.14	0.10
1 Jan 89	23	CAM	0.07	0.05	0.09	0.02	0.10
4 Jan 89	20	CAM	0.19	0.09	0.09	0.02	0.28
6 Jan 89	30	CAM	0.06	0.05	0.06	0.02	0.05

Table 12c. The Ratio of the Root-Mean-Square Errors of 300-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 300-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
1 Jul 88	85	EUR	0.14	0.14	0.16	0.19	0.11
1 Jan 89	82	EUR	0.21	0.10	0.12	0.09	0.08
4 Jan 89	100	EUR	0.17	0.10	0.19	0.12	0.11
6 Jan 89	96	EUR	0.35	0.13	0.12	0.13	0.08
1 Jan 89	23	CAM	0.06	0.08	0.01	0.02	0.04
4 Jan 89	20	CAM	0.07	0.04	0.04	0.02	0.06
6 Jan 89	30	CAM	0.14	0.05	0.01	0.01	0.06

Table 13a. The Ratio of the Root-Mean-Square Errors of 850-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 850-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
2 Jul 88	94	EUR	0.53	0.21	0.13	0.13	0.10
2 Jan 89	102	EUR	0.21	0.29	0.16	0.06	0.11
5 Jan 89	96	EUR	0.18	0.18	0.17	0.08	0.12
7 Jan 89	91	EUR	0.23	0.38	0.29	0.16	0.12
1 Jul 88	25	CAM	0.08	0.05	0.06	0.05	0.07
2 Jan 89	27	CAM	0.09	0.11	0.11	0.02	0.06
5 Jan 89	20	CAM	0.13	0.06	0.11	0.06	0.09
7 Jan 89	26	CAM	0.23	0.10	0.03	0.05	0.04

Table 13b. The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the RMS Errors of HIRAS 500-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
2 Jul 88	94	EUR	0.27	0.32	0.27	0.11	0.12
2 Jan 89	101	EUR	0.28	0.16	0.21	0.11	0.10
5 Jan 89	97	EUR	0.30	0.23	0.16	0.07	0.01
7 Jan 89	96	EUR	0.24	0.22	0.14	0.11	0.17
2 Jul 88	25	CAM	0.03	0.07	0.07	0.04	0.04
2 Jan 89	29	CAM	0.13	0.09	0.06	0.02	0.06
5 Jan 89	20	CAM	0.12	0.10	0.10	0.01	0.07
7 Jan 89	29	CAM	0.29	0.03	0.03	0.03	0.07

Table 13c. The Ratio of the Root-Mean Square Errors of 300-mb RAP Analyses Derived Using 36-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of HIRAS 300-mb Analyses

Date	Samples	Region	Z_R/Z_H	T_R/T_H	Td_R/Td_H	U_R/U_H	V_R/V_H
2 Jul 88	89	EUR	0.31	0.11	0.13	0.06	0.07
2 Jan 89	96	EUR	0.23	0.09	0.19	0.07	0.16
5 Jan 89	92	EUR	0.25	0.07	0.23	0.09	0.11
7 Jan 89	84	EUR	0.41	0.09	0.19	0.09	0.07
2 Jul 88	25	CAM	0.06	0.06	0.04	0.03	0.05
2 Jan 89	28	CAM	0.12	0.02	0.07	0.02	0.06
5 Jan 89	19	CAM	0.08	0.03	0.04	0.02	0.04
7 Jan 89	22	CAM	0.16	0.03	0.06	0.02	0.05

The RAP analyses derived using the 12-hr forecast error correlation models have much smaller errors than the HIRAS analyses, especially above 850 mb and particularly for wind and humidity. Tables 11a - 11c show, however, that wind and humidity have relatively larger analysis errors than geopotential or temperature; therefore, HIRAS is probably weak in analyzing wind and humidity.

Similarly, Tables 12a, 12b, and 12c show that the RAP analyses derived from the 36-hr forecast error correlation models are better than HIRAS analyses, although not as overwhelmingly in Eurasia compared to RAP analyses derived from the 12-hr forecast error correlation models. In the central American region, however, almost all RAP errors are an order of magnitude smaller than HIRAS errors. No matter how much caution is used when interpreting these results, the inescapable conclusion is that RAP performs superior analyses.

An interesting comparison of the 12-hr forecast error correlation models to the 36-hr models can be obtained by considering the ratio of their resulting analysis errors, as shown in Table 14. The preferred model is, as expected, the 12-hr model.

Table 14. The Ratio of the Root-Mean-Square Errors of 500-mb RAP Analyses Derived Using 12-Hr Forecast Error Correlation Models to the Root-Mean-Square Errors of RAP 500-mb Analyses Derived Using 36-Hr Forecast Error Correlation Models

Date	Region	Z_{12}/Z_{36}	T_{12}/T_{36}	Td_{36}/Td_{36}	U_{12}/U_{36}	V_{12}/V_{36}
2 Jul 88	EUR	0.41	0.69	0.67	0.91	0.75
2 Jan 89	EUR	0.75	0.69	0.67	0.64	0.80
5 Jan 89	EUR	0.60	0.43	0.44	1.0	0.10
7 Jan 89	EUR	1.25	0.55	0.86	0.27	0.59
2 Jan 89	CAM	0.54	0.56	3.0	1.0	1.67
5 Jan 89	CAM	1.58	0.90	0.90	2.0	4.0
7 Jan 89	CAM	0.21	0.66	3.0	0.67	0.71

2.4 STC TASK 4

The OSSEs required the development of several global observation databases described in the interim report. The "perfect" observation (BASELINE) database was developed by interpolating the ECMWF T-106 forecasts at 6-hr intervals to the FGGE-2b observation points from 19 January - 1 February 1979. The simulated observation (CONTROL) database was developed by inserting appropriate errors into the perfect observations. The TACOBS (battlefield) database was developed by denying data to the eastern portions of the RAP regions.

The complete set of OSSEs would have consisted of 2 regions x 3 unique cycles x 9 types of first guesses (12-hr, 18-hr, 24-hr, and 36-hr forecasts; 12- and 24-hr-old RAP analyses; and 3 mismatched RWM forecasts with OLDOBS [a 12-hr forecast valid at time T with offtime observations at time T-6, an 18-hr RWM forecast valid at time T+6 with offtime observations at time T, and a 24-hr forecast valid at time T with offtime observations at T-6] x 4 observation scenarios (BASELINE, CONTROL, ALLOBS, and TACOBS), a total of 216 RAP analyses (reduced to 144 with the late failure of the 18-hr forecast error correlation models and the lack of the ALLOBS³ scenario). Even with fewer cases to consider, insufficient time remained to perform and analyze all of them.

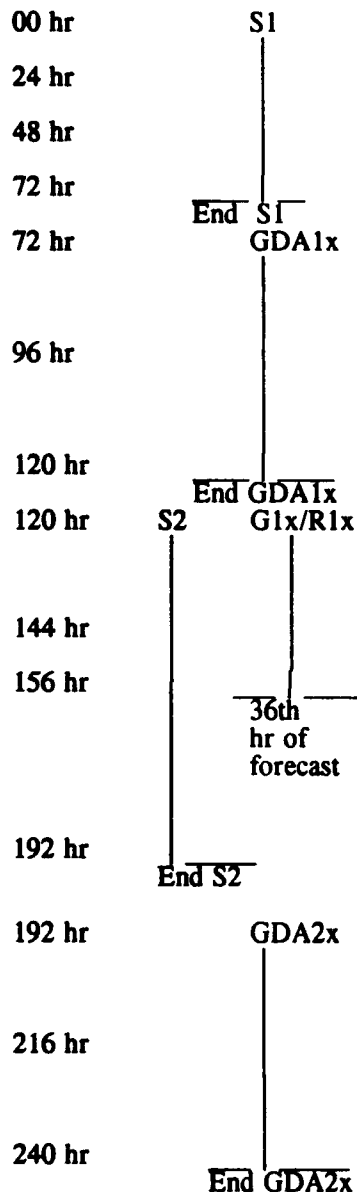
When production had to be stopped shortly after the end of the contract's performance period because insufficient funds remained in the computer account, OSSEs were completed for all three cycles (the first was initialized at T-106 hour 00, 16 January 1979, 00 UTC); the second at T-106 hour 120, 21 January 1979, 00 UTC; and the third at T-106 hour 240, 26 January 1979, 00 UTC, respectively) and each of the three observation scenarios (as shown in Fig. 9) for the 12-, 24-, and 36-hr forecasts. In the Eurasian region, which is almost all land, RAP analyses that included satellite data were virtually identical to analyses that excluded satellite data, as expected because no satellite data could be simulated over land. Over the central American region, however, two sets of analyses were performed, one with and the other without satellite data, to determine the effect of incorporating satellite data into RAP analyses.

Consequently, the OSSEs consisted of 81 unique experiments: (1 Eurasian region x 3 cycles x 3 first guesses x 3 observation scenarios) + (1 central American region x 3 cycles x 3 first guesses

³ The ALLOBS scenario was also requested but could not be developed. There was insufficient time to simulate observations from the proposed systems (namely, a MARK IV-B van), which when added to the existing (CONTROL) data would have defined the ALLOBS set.

T-106
Forecast
Hr

Input/Output Required for the OSSEs



Run the PL/GP Global Spectral Model (GSM) for 72 hrs to generate a "spin-up" atmosphere (S). Use the T-106 forecast at 00 hr to initialize the GSM forecast. This 72-hr forecast is S1, which is the first guess for the PL/GP Statistical Analysis Program (ASAP).

Three 48-hr global data assimilations (GDA), using ASAP and the GSM (for a first guess), were required at 6-hr intervals. x=1 is the BASELINE scenario, x=2 is the simulated (i.e. CONTROL) scenario, and x=3 is the TACOBS scenario.

Clearly, the GDAs generated are scenario dependent. The observations for x=2 and 3 are simulated (by interpolating the T-106 forecasts to FGGE-2b observation points and inserting observation errors) FGGE-2b observations. The GDA begins on 19 Jan/00 UTC.

Run a second 72-hr GSM spin-up forecast (S2), starting at 21 Jan 79/00 UTC. Also, run 36-hr GSMs (G1x) to provide boundary conditions to run the RWMs (R1x), using the 48th hr of above GDAs as initial conditions. The RWM forecasts (R1x) are R11, R12, and R13, where x is as described above.

Store the 12-, 18-, 24-, and the 36-hr RWM forecasts, which will be the first guesses for RAP analyses.

Use S2 to initialize a second run of four 48-hr GDAs at 6-hr intervals. The GSM and ASAP combination using: perfect observations (BASELINE) is GDA21, existing observations (CONTROL) is GDA22, and the specific data-denied scenario (TACOBS) is GDA23.

As above, the simulated FGGE-2b observations (T-106 forecasts interpolated to FGGE-2b observation points and distorted by adding observation errors) are used. These GDAs begin at T-106 hour 192 (24 Jan 79, 00 UTC) and end on hour 240 (26 Jan 79, 00 UTC).

(continued)

Figure 9. A depiction of the process for generating the OSSE databases, from 16 January 1979 at 00 UTC, which is the 00 hr for the T-106 forecasts.

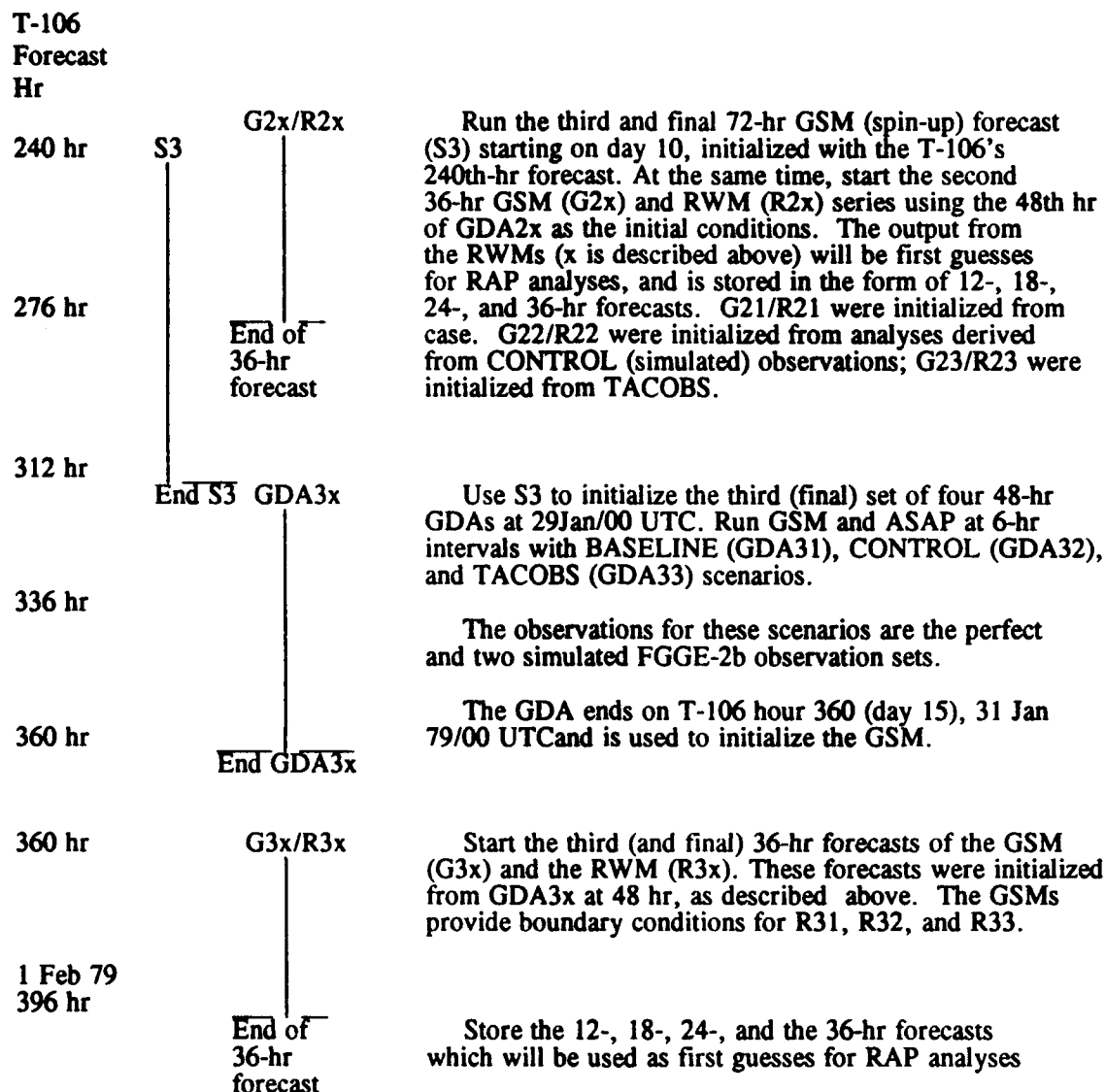


Figure 9. A depiction of the process for generating the OSSE databases from 16 January 1979 at 00 UTC to 1 February 1979 at 1200 UTC.

x 6 observation scenarios [= 3 observation sets including satellite data + 3 observation sets excluding satellite data]) = $3 \times 3 \times 3 + 3 \times 3 \times 6 = 81$ OSSEs.

2.4.1 Development of the OSSE Databases

The following is a brief, updated summary of the detailed description provided in Section 2.4 of the RAP Interim Report. (See Fig. 9 for a depiction of the development of the OSSE databases.) The

OSSEs began with a 72-hr GSM (spin-up) forecasts initialized from the FGGE-2b perfect observation network at 16 January 1979, 00 UTC; 21 January 1979, 00 UTC; and 26 January 1979, 00 UTC (T-106 forecast hours 00, 120, and 240, respectively). As discussed in detail below, the OSSEs consist of three scenarios developed by assimilating different (perfect, simulated, and TACOBS) First Global Atmospheric Research Program (GARP) Global Experiment (FGGE)-2b observation sets at 6-hr intervals for 48 hr.

The 72-hr GSM forecast was initialized from a 00-hr GSM field derived from preprocessed T-106 forecasts at hours 00, 120, and 240. Each of these spin-up atmospheres was the first guess for the AFGL Statistical Analysis Package (ASAP), whose analyses initialized the GSM for the global data assimilation (GDA).

Each GDA started with the 72-hr GSM forecast as a first guess, and one of the observation scenario data sets as observations yielding an ASAP-produced analysis. These analyses are used to initialize 6-hr GSM forecasts, which become the first guesses for ASAP to analyze the next set of observations. This data assimilation process continued at 6-hr intervals for seven additional cycles, completing the 48-hr GDAs. The 48-hr GDAs were therefore processed for three observation scenarios (BASELINE, CONTROL, and TACOBS with no satellite data) in Eurasia and six observation scenarios (the preceding three with satellite data and three without satellite data).

A complete analysis of the three scenarios is provided in Section 2.4.2. Briefly, in the BASELINE database the GDA apparently compromised between a first-guess forecast and the observations, that is, the observations brought the first guess about halfway towards "truth" (the verifying T-106 forecast). The CONTROL database appeared to be similar to the BASELINE database through the 48-hr GDA except for the wind fields.

When plots of the 48th-hr of the BASELINE GDA were overlaid on the plots of the 48th-hr of the CONTROL GDA, the contours matched nearly one-for-one with occasional offsets, which could not be detected in a side-by-side comparison (see Figs. 10a, 10b, 10c, 11a, 11b, 11c, 12a, 12b, and 12c). An analysis of the RMS differences between any two variables, however, illustrated that the insertion of errors did have an effect, even if not visually obvious in the RAP windows. The TACOBS GDA, even with data removed from the eastern portion of the Eurasian region and the southwestern portion of the

central American region, was surprisingly similar to the CONTROL scenario in terms of the visual comparisons (see Figs. 10a, 10b, 10c, 11a, 11b, 11c, 12a, 12b, and 12c).

The 48th hr of the GDA of the differing observation scenarios was used to initialize the GSM, which was run for 36 hr to generate forecasts at 6-hr intervals that provided the boundary conditions for the RWM. The RWM was run for 36 hr to produce the 12-, 18-, 24- and 36-hr forecasts, first guesses for RAP.

RAP analyses were performed at 12, 24, and 36 hr following the T-106 forecast hour of 120 (21 January), 240 (26 January), and 360 (31 January), respectively, for each of the observation scenarios, as shown in Fig. 9 and discussed in Section 2.4. Each of the RAP analyses used appropriate observation sets to ensure consistency with the GDAs. Separate analyses, one with satellite data and another without satellite data, were performed over the central American region (no satellite data were available over the Eurasian landmass). All analyses and the RWM forecast were compared to truth (appropriate T-106 forecasts) to determine analysis errors.

Sections 2.4.2, 2.4.3, and 2.4.4 describe an analysis of the BASELINE (perfect), CONTROL (simulated), and TACOBS (simulated data denial) OSSE GDA fields. Section 2.4.5 describes the RWM forecasts initialized from the 48th hr of the three GDAs. Section 2.4.6 shows tables of the RWM forecast errors and the RAP analysis errors at 500 mb. An analysis of these errors is the purpose for conducting the OSSEs. Similar tables of errors at 850 and 500 mb are provided in Appendix A.

2.4.2. An Analysis of the BASELINE OSSE Global Data Assimilation Fields

Each of the following two tables provide objective comparisons, through rms differences, of the OSSEs. Table 15 shows the results of the BASELINE OSSE, which has three distinct cycles, separated by slants (/) in the tables, initialized at T-106 hour 00 (16 January 1979, 00 UTC), T-106 hour 120 (21 January 1979, 00 UTC), and at T-106 hour 240 (26 January 1979, 00 UTC), respectively. The tables have six sets of differences that were calculated along two representative latitude circles (29° N and 49° N): three following the 72-hr GSM forecast but before the 48-hr global data assimilation (GDA) begins, and three after the GDA.

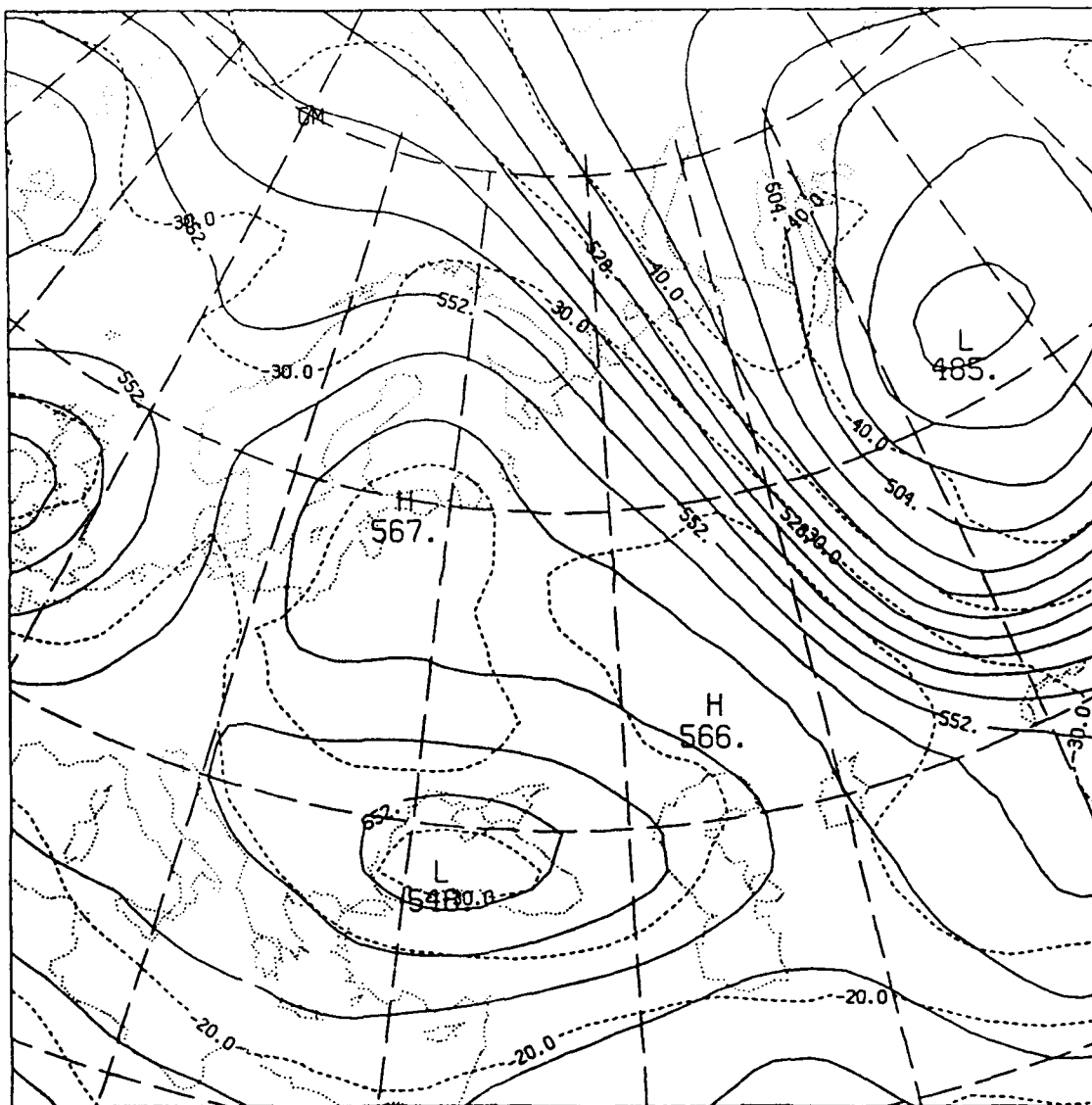


Figure 10a. Simulated GDA 48-hr analysis of 500-mb heights (dm) and temperature (C) valid 21 January 1979, 00 UTC over Eurasia for BASELINE.

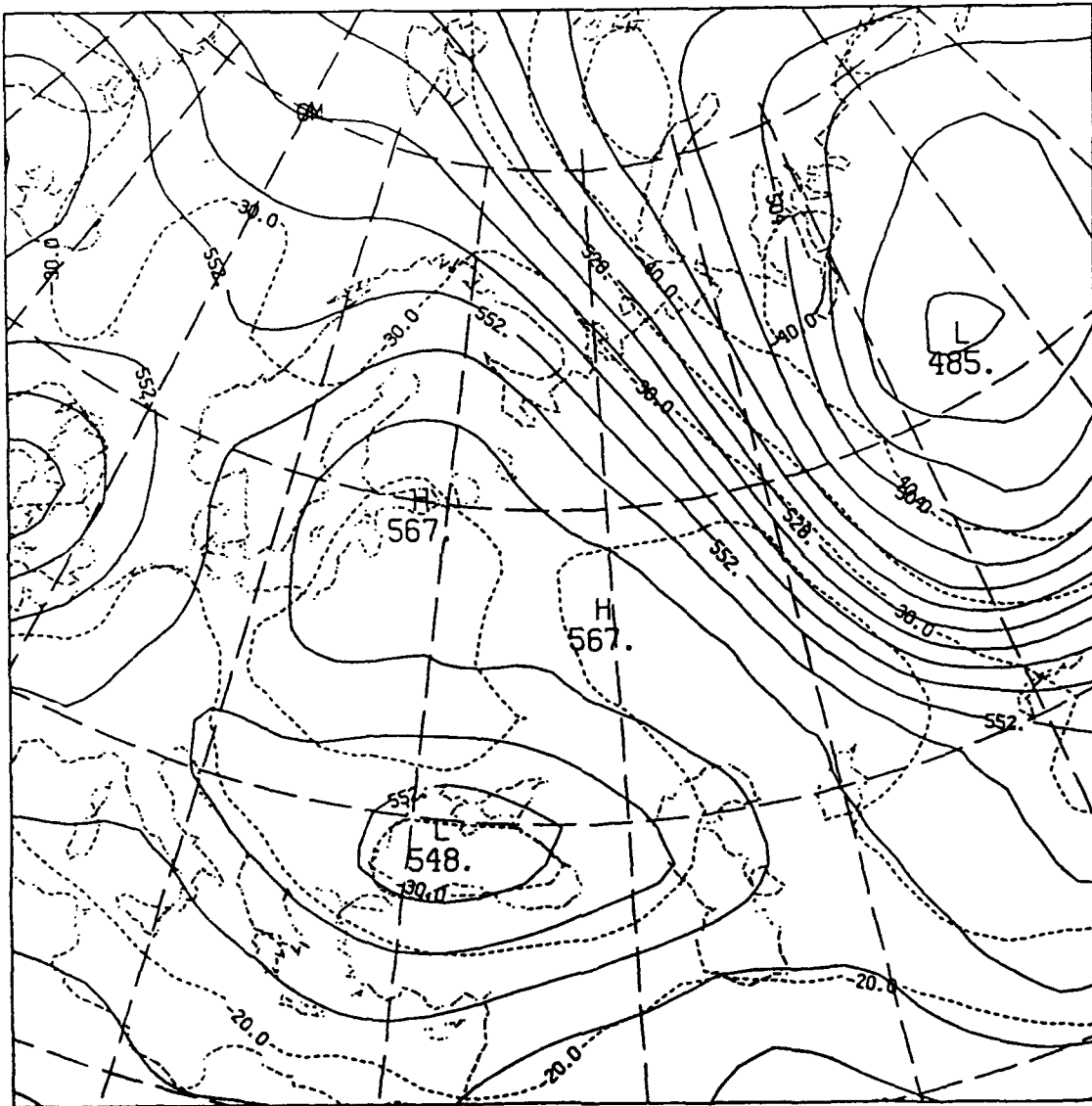


Figure 10b. Same as Figure 10a except for CONTROL.

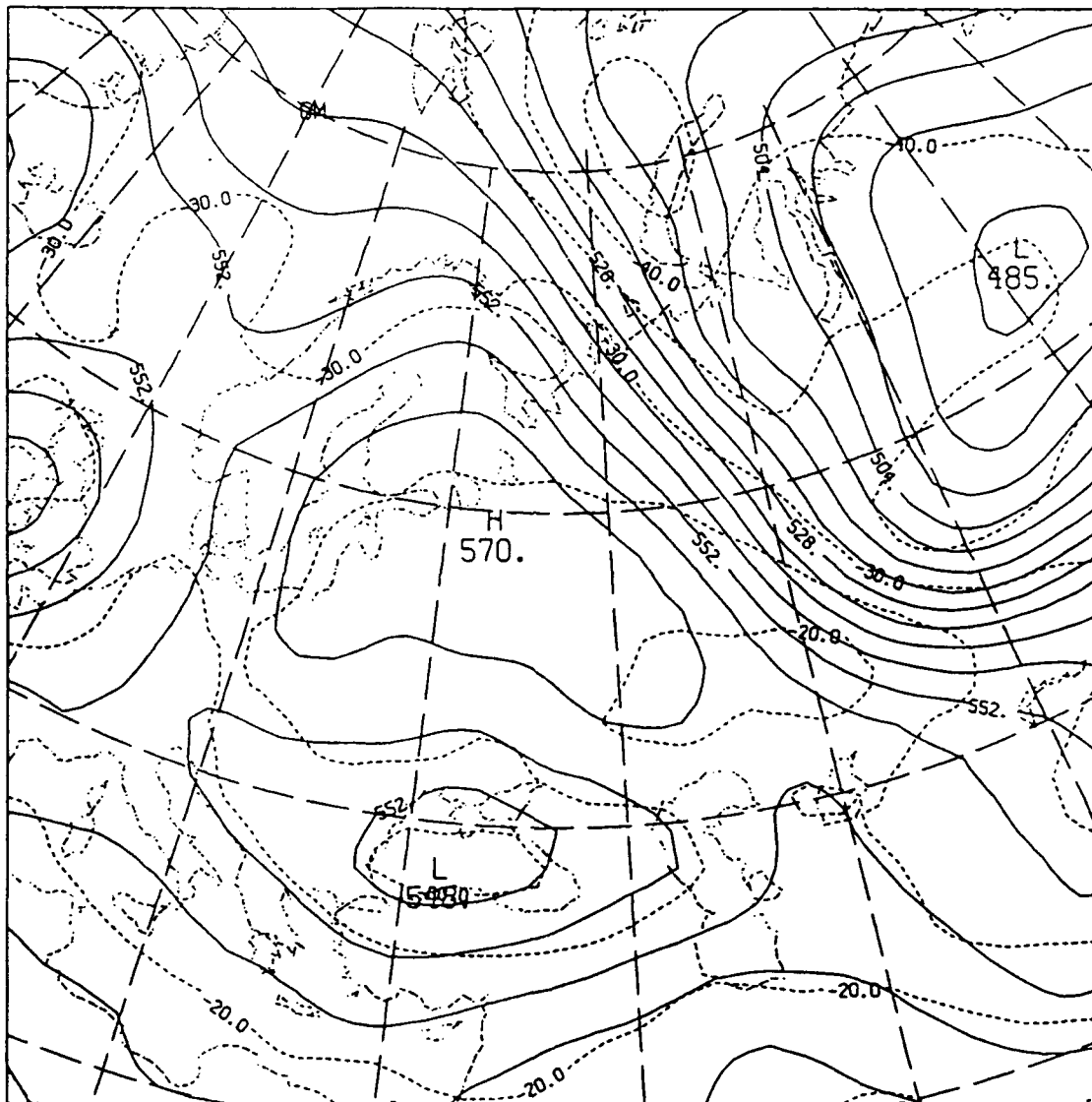


Figure 10c. Same as Figure 10a except for TACOBS.

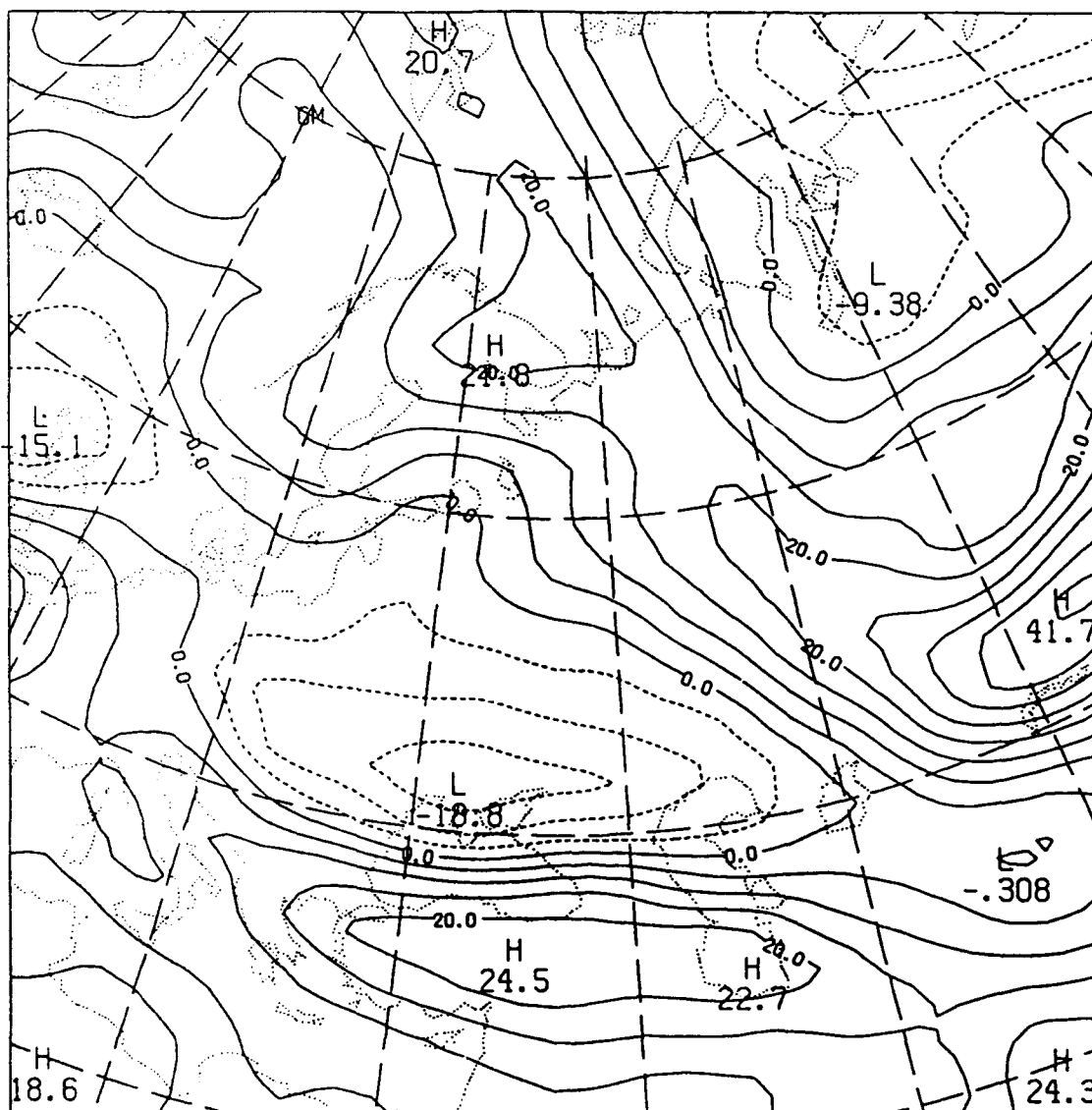


Figure 11a. Simulated GSA 48-hr analysis of 500-mb u-wind (m/s) valid 21 January 1979, 00 UTC over Eurasia for BASELINE.

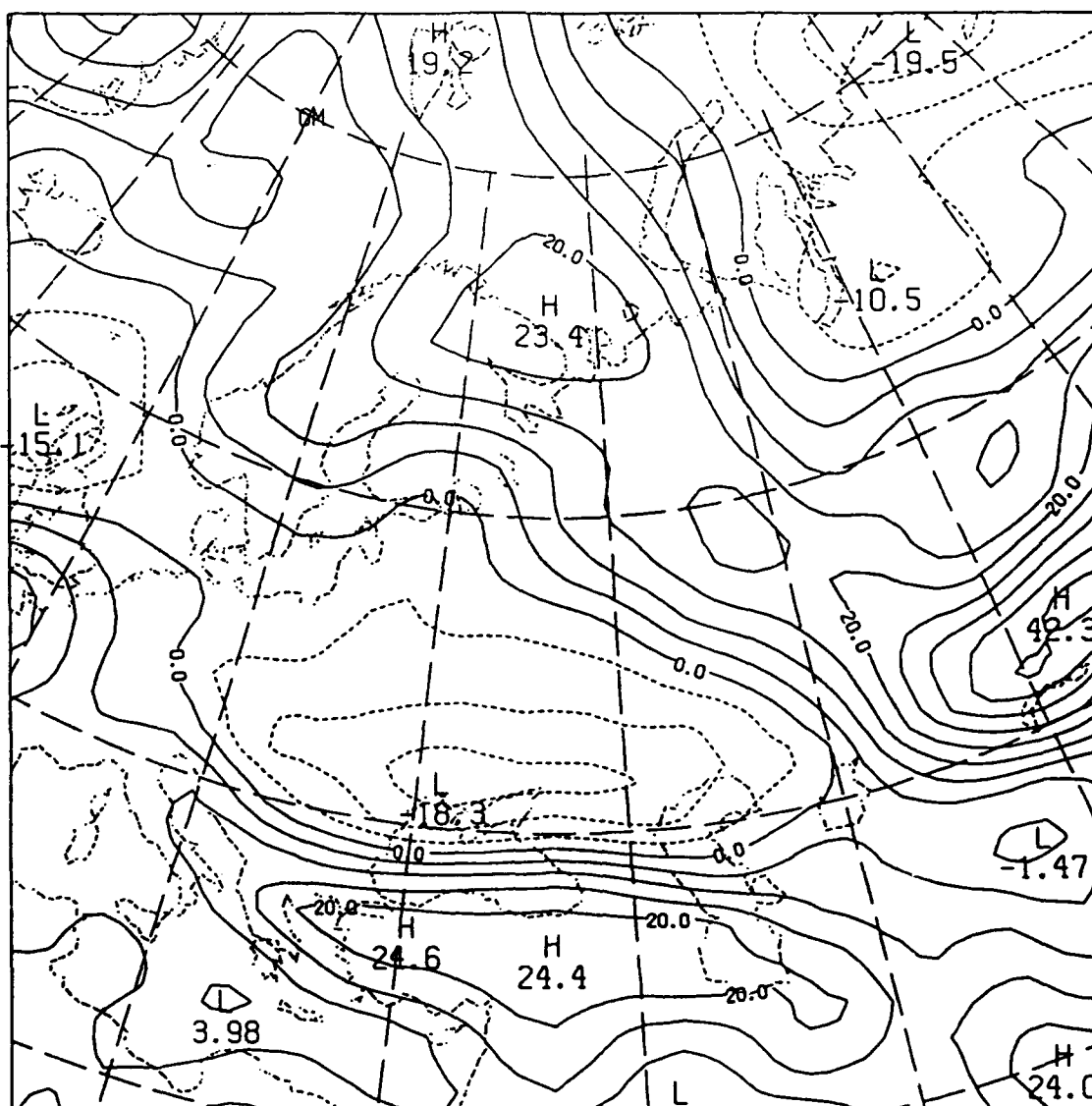


Figure 11b. Same as Figure 11a except for CONTROL.

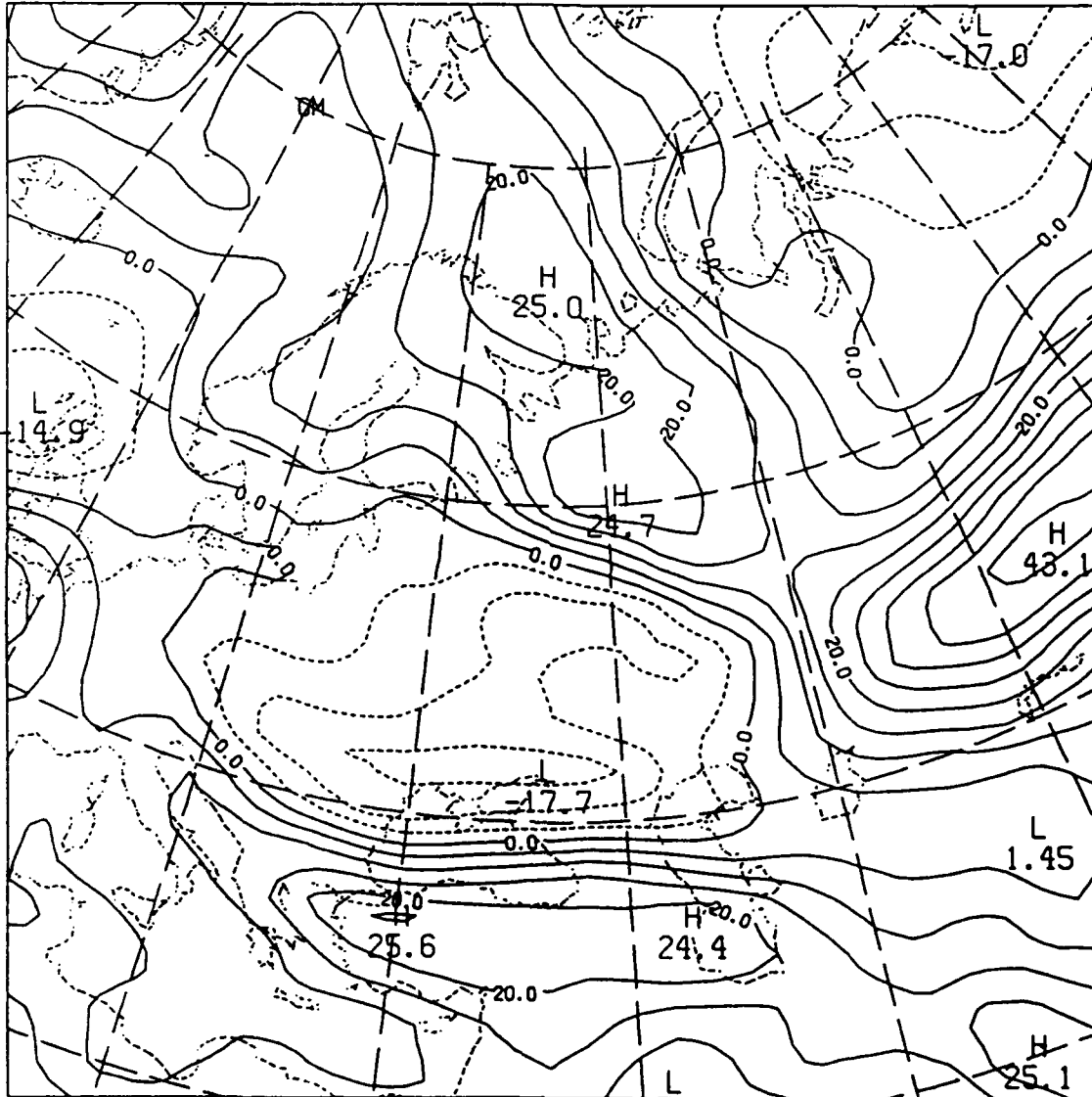


Figure 11c. Same as Figure 11a except for TACOBS.

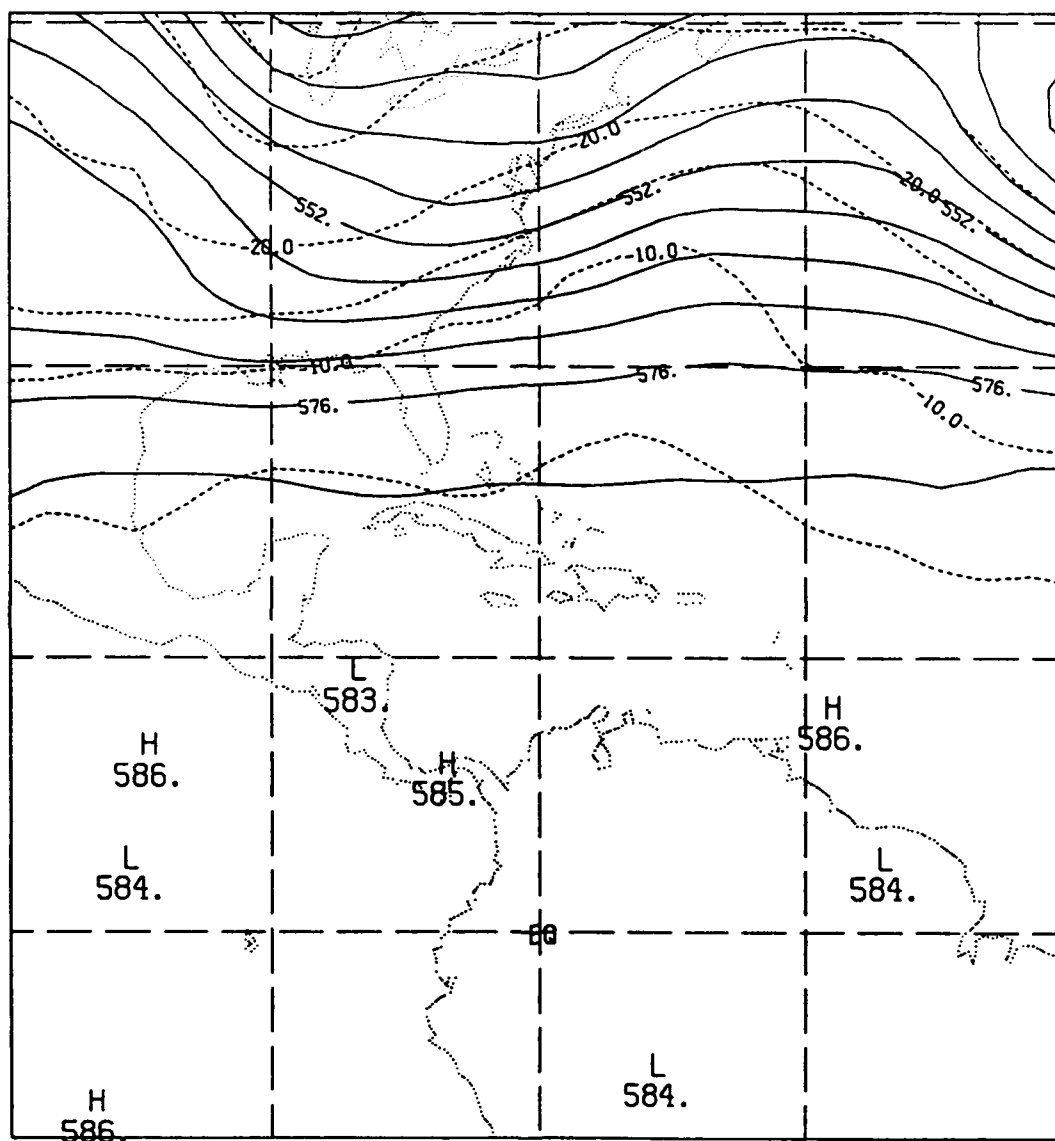


Figure 12a. Simulated GDA 48-hr analysis of 500-mb heights (dm) and temperature (C) valid 21 January 1979, 00 UTC over central America for BASELINE.

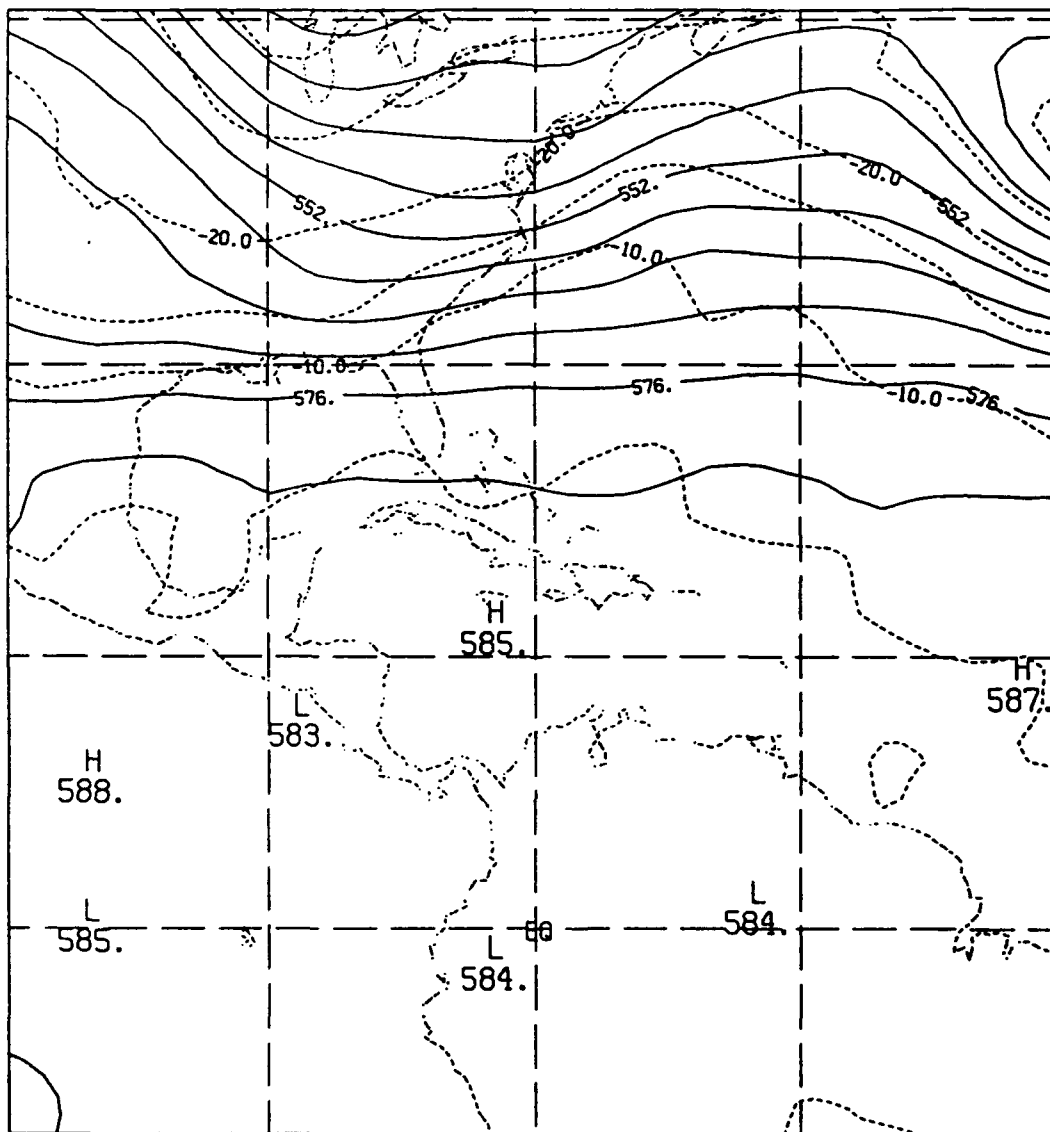


Figure 12b. Same as Figure 12a except for CONTROL.

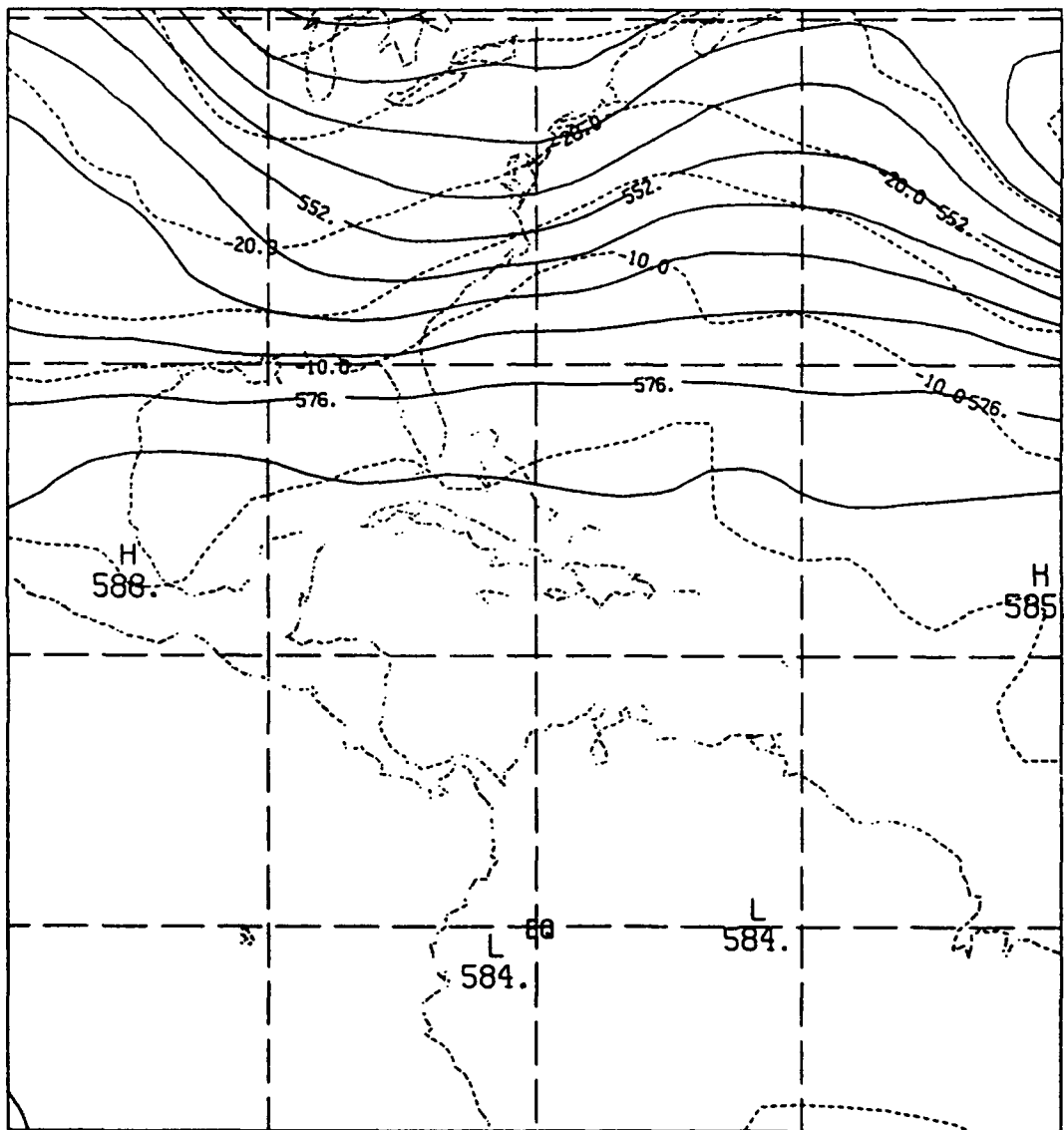


Figure 12c. Same as Figure 12a except for TACOBS.

The differences between forecasts and analyses [of temperature (T), u- and v-wind components (u and v), and relative humidity (RH)] are calculated at approximately 500 mb along the two latitude circles of relevance to the RAP regions. Also, plots (not shown) of these fields are discussed to highlight differences or note similarities in the RAP windows. Specifically, the differences shown in the tables are between (1) "truth" (the ECMWF T-106 Nature Run forecasts) and the 72-hr GSM spin-up forecasts, (2) truth and an analysis using the 72-hr forecast as a first guess, (3) the analysis and the first guess, (4) truth and the 6-hr GSM forecast initialized from the 42nd hr of the GDA, (5) truth and the 48th hr of a GDA, and (6) the 48th hr of the GDA and the 48-hr GSM forecast.

The first set of comparisons (the BASELINE experiment) is presented in greater detail to establish a solid foundation for the OSSEs. All observations are "perfect"; therefore, all errors are due to forecasts or analysis procedures. The second set of comparisons (the CONTROL experiment) begins with the same 72-hr GSM spin-up forecast, but only simulated observations are assimilated. Thus, assessing the effect of including observation errors is a straightforward task.

As expected, even though each BASELINE experiment is separated by 5 days, Table 15 shows that each has similar characteristics. There are differences, especially the 72-hr GSM forecasts of winds, but the overall performance of the forecast/analysis system (GSM/ASAP) is more distinguished by the similarity among, rather than differences between, the experiments. In general, the forecast errors are larger than the analysis errors, which in turn are larger than the differences between the analysis and the first guess.

The results of the analysis shown in Table 15 (and similarly Tables 16-19) must be seen in the light of the fact that the differences are taken at intervals of 5° of longitude along the 29° and 49° latitude circles. Therefore, much of the data is sampled over the data-void oceans, where the first guess will dominate the analysis; hence, the analysis may not on average be pulled closer to the truth. Also, temperature observations were apparently assimilated better than wind or humidity; nevertheless, data assimilation improved all analyses, although the humidity analysis only slightly. The comparatively ineffective assimilation of humidity may be due to the relatively fewer BASELINE humidity observations, which are available only from ROABs, that could be assimilated.

Table 15. The Root-Mean-Square Differences at Approximately 500 mb from the Three BASELINE OSSEs (shown separated by a slant) that Began at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979

Latitude/ Element	Forecast Error (Truth at 72 hr vs. 72-hr Spin-Up)	Analysis Error (Truth vs. the Analysis at 72 hr)	Analysis vs. Spin-Up at 72 hr
29° N			
T (K/10)	22/21/21	14/12/12	18/17/16
U (m s ⁻¹ /10)	48/50/53	43/40/40	18/22/25
V (m s ⁻¹ /10)	51/73/45	48/52/41	21/26/23
RH (%)	24/25/26	20/21/24	12/11/14
49° N			
T (K/10)	38/38/32	17/25/17	29/24/24
U (m s ⁻¹ /10)	72/53/86	48/40/57	41/30/42
V (m s ⁻¹ /10)	78/70/64	57/52/41	39/32/36
RH (%)	41/25/29	36/21/23	17/16/16
	Truth vs. Forecast at 120 hr (6-h GSM Forecast from the 42nd hr of the GDA)	Truth vs. Analysis at 120 hr (48th hr of the 48-hr GDA)	Analysis vs. Forecast at 120 hr
29° N			
T (K/10)	11/13/12	10/11/10	7/8/7
U (m s ⁻¹ /10)	38/40/42	37/35/31	11/18/23
V (m s ⁻¹ /10)	35/38/38	34/32/41	12/25/27
RH (%)	19/21/25	19/20/22	8/10/12
49° N			
T (K/10)	11/11/13	9/9/9	8/10/10
U (m s ⁻¹ /10)	32/38/52	32/32/44	15/15/16
V (m s ⁻¹ /10)	38/33/51	30/31/46	16/13/21
RH (%)	22/26/24	18/24/22	12/13/12

Note that, except occasionally for humidity, the values (errors) in the upper half of the table are larger than the values in the lower half. This distribution is expected because the first guess forecast is 72 hr old in the upper half of the table, while the first guess is 6 hr old in the lower half. The 6-hr forecast is expected to be the better; the errors of the GDA should be smaller given the better first guess.

Visual comparison of plots of truth, 72-hr spin-up forecast, and ASAP BASELINE analyses at 500 mb at T-106 hour 72 yielded the subjective observations immediately below:

Central American Region

- Z:** The forecast trough should have been further off the New England coast and deeper. The analyzed trough was a compromise between truth and the forecast.
- U:** The forecast jet off the mid-Atlantic coast was too weak by 7 (forecast 40 vs. true 47) m/s, but the forecast maximum over Texas was too strong (30 vs. 27). The analysis splits the difference between truth and the forecast.
- V:** The forecast was slightly weaker over NY (-40.0 vs. -42.6) and considerably weaker over the Central Plains (15.6 vs. 20.2).
- RH:** The forecast was too dry over the northeastern United States, too moist over Brazil, and too moist over the northeastern corner of the region. The analysis was only a slight improvement of the forecast.

Eurasian Region

- Z:** The forecast low near the pole was deeper (484 vs. 490 dm), and the forecast high over Scandinavia was weaker (567 vs. 571 dm) than truth. The forecast was excellent over the Balkans.

- U:** Forecast jet over the Barents Sea was too strong (57.3 vs. 50.8 m/s); so was the maxima over the Mediterranean Sea (34 vs. 30). The maxima in central Siberia is displaced slightly too far. The analysis did not bring the forecast as close to truth as it typically had.
- V:** The forecast jet over Turkey was too strong (34.5 vs. 28.2 m/s), and the maxima over Poland was too weak (18.1 vs. 24.8).
- RH:** The forecast features verified quite well everywhere except over the Mediterranean Sea, where the forecast was much too moist, and over Scandinavia, where it was much too dry. The analysis was a slight improvement most everywhere, especially over Scandinavia.

Visual comparison of plots of truth, 6-hr forecast initialized from 42nd hr of the BASELINE global data assimilation, and ASAP analyses at 500 mb at the T-106 hr 120 resulted in the following subjective analysis:

Central American Region

- Z:** The dominant ridge-trough was forecast about 150 km east of truth. The analysis improved the position of the ridge but slightly weakened it.
- U:** The forecast jet core was weakened about 2 m/s and shifted about 200 km to the south. The analyzed position of the jet core was correct but 1.3 m/s too weak.
- V:** The forecast maxima off the east coast of the United States was weakened by 6 m/s; the analyzed maxima was too weak by 3.6 m/s.
- RH:** The forecast and analysis were approximately the same and close to truth.

Eurasian Region

- Z:** The forecast low over central Siberia was 30 m too deep. The analysis was within 10 m of truth.

- U:** Small forecast differences were usually within 1 m/s. The analysis was a match of truth.
- V:** There were only slight forecast differences of about 1 m/s. The analysis was in error over central Siberia by 1.2 m/s.
- RH:** Both the forecast and analysis nearly matched the truth.

Visual comparison of plots of truth, the 72-hr spin-up GSM forecast (initialized at T-106 hour 120), and 500-mb BASELINE analyses valid at T-106 hour 192 resulted in the following subjective analysis and conclusions:

Central American Region

- Z:** Forecast low over the Central Plains was 60 m too deep. Forecast low east of Cape Cod was good. The analysis basically splits the difference between truth and the forecast.
- U:** A good forecast that the analysis could improve only slightly.
- V:** The forecast maxima over the Central Plains was 17 m/s too strong; the maxima east of Cape Cod was 7 m/s too strong. The analysis retained much of the error (12 m/s) over the Central Plains; however, the analyzed maxima east of Cape Cod reduced the error to only 1 m/s.
- RH:** A good forecast except 1500 km east of Cape Cod, where the RH was too high. The analysis was slightly better there.

Eurasian Region

- Z:** The low north of Sweden should have been forecast 80 m deeper. The low in western Turkey should have been 100 m deeper. The analysis approximately split the difference between truth and the forecast.

- U:** A good forecast except over and just east of the Caspian Sea, where the forecast maxima was almost 5 m/s too weak. The analysis made little change.
- V:** The maxima forecast over Lithuania was 5 m/s too weak. The analysis either eliminated or substantially reduced most errors.
- RH:** The forecast over the northeastern part of Eurasia was poor; over the remainder of the region forecast was good. The analysis reduced most of the errors.

Visual comparison of plots of truth, 6-h forecast initialized from 42nd hr of BASELINE global data assimilation, and ASAP analyses at 500 mb at the T-106 hour 240 resulted in the following subjective analysis and conclusions:

Central American Region

- Z:** The forecast position of major features was slightly too far east and weak.
- U:** The forecast strong jet from the Central Plains to South Carolina and eastward was well positioned but weakened by 3 m/s off the southeastern coast of the United States and eastward. The analysis hardly changed the forecast.
- V:** The forecast maxima east of New Jersey was about 200 km too far south, and the maxima over the Ohio Valley was about 200 km too far east; the analysis did not provide a correction.
- RH:** The forecast and analysis were approximately the same and close to truth.

Eurasian Region

- Z:** The forecast main features were very good except for the too-weak short wave over and north of the Black Sea. The analysis, however, corrected the error.

- U:** Small forecast differences were usually within 1 m/s. The analysis was slightly further from truth than the forecast!
- V:** Only slight forecast differences of about 1 m/s; all major features were forecast. The analysis nearly corrected the small errors.
- RH:** Both the forecast and analysis nearly matched truth.

Visual comparison of plots of truth, 72-h spin-up forecast, and BASELINE ASAP analyses at 500 mb at T-106 hour 312 resulted in the following subjective analysis and conclusions:

Central American Region

- Z:** The forecast east coast ridge should have been 600 km inland, and the low 1500 km east of New Jersey was about 300 km too far south-southeast. The analyzed low was 30 m too deep but positioned correctly. The ridge was positioned nearly correctly but not sharp enough.
- U:** The forecast jet core off the mid-Atlantic coast was too strong by 2 (forecast 43 vs. true 41) m/s and 300 km too far southeast, and the forecast maxima over Texas was too strong (38 vs. 45) and 200 km too far north. The analysis hardly affected forecast.
- V:** The forecast over the central United States was bad, missing the Central Plains maxima and southerly jet over Illinois. The northerly jet 700 km east of Delaware is forecast 2 m/s too strong and 300 km to the southeast. The analysis improved the forecast, especially over Illinois.
- RH:** The forecast was too dry over the midwestern United States, and too moist off North Carolina. The moist and dry areas were forecast about 700 km too far eastward. The analysis barely improved the forecast.

Eurasian Region

- Z:** There were three main forecast features: a low just west of Sweden with a trough extending southeast to the Black Sea, a high north of the Caspian Sea and ridge extending to the north-northwest, and a low about 2000 km to the east-northeast of the high. All were forecast to be too strong, ranging from 30 to 80 m, and positioned about 300 km too far south. The analysis corrected most of the differences.
- U:** The forecast jet cores were relatively weak; errors were less than 5 m/s. The analysis brought the forecast close to truth.
- V:** The forecast jet core north of the United Kingdom was excellent, the southerly jet from the Baltic Sea to the Black Sea was displaced to the east-southeast, but the northerly jet over central Eurasia was forecast well. The analysis provided little help to the forecast.
- RH:** The main forecast features verified quite well everywhere except some were displaced. The analysis was only a slight improvement.

Visual comparison of plots of truth, 6-h forecast initialized from 42nd hr of BASELINE global data assimilation, and ASAP analyses at 500 mb at T-106 hour 360 resulted in the following subjective analysis and conclusions:

Central American Region

- Z:** The forecast low 2000 km east of Cape Cod was positioned well but 50 m too weak. The analysis had no effect.
- U:** The forecast strong jet approximately 35°N was well positioned and only slightly weakened. The analysis hardly changed the forecast.

V: The forecast maxima 1700 km east of New Jersey was positioned well but 4.5 m/s too weak; the maxima over the Central Plains was also positioned well but was 8 m/s too strong. The analysis did not provide a correction.

RH: The forecast and analysis were similar except over the western Gulf of Mexico, where they are drier than truth.

Eurasian Region

Z: The forecast and analyzed main features were very close to truth.

U: The weak jet core west of Sweden was forecast 4 m/s too weak; the analysis error was 2 m/s. Everywhere else the analysis and forecast nearly matched truth.

V: Only slight forecast differences of about 1 m/s. The maxima over Austria was analyzed 3 m/s too strong, that is, worse than the first guess.

RH: Both the forecast and analysis nearly matched truth.

2.4.3. An Analysis of the CONTROL OSSE Global Data Assimilation Fields

Table 16 shows the three CONTROL OSSEs, which assimilated simulated observations (that is, appropriate errors were added to the "perfect" observation sets) during analysis/forecast system cycles starting from an analysis based on the 72-hr GSM forecast through the 48-hr GDA. Even though each of the three experiments is separated by 5 days, the table has similar characteristics but exhibits some notable differences too. The discussion below is an analysis of the CONTROL OSSE from Table 16.

Similar to the BASELINE scenario, the differences at the completion of the 48-hr GDAs (lower half of the table) are smaller in general (the zonal wind and humidity being the exception along 29°N) than the differences following the 72-hr GSM forecasts. Except for the zonal wind, whose simulated errors appear to be large, these results are as expected. Strangely enough, however, the assimilation of humidity observations with errors somehow brought the analyses closer to truth at higher latitudes.

Table 16. The Root-Mean-Square Differences at Approximately 500 mb from the Three CONTROL OSSEs (shown separated by a slant) that Began at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979 (each OSSE lasted for 120 hr, consisting of a 72-hr GSM forecast to begin a 48-hr global data assimilation)

Latitude/ Element	Forecast Error (Truth at 72 hr vs. 72-hr Spin-Up)	Analysis Error (Truth vs. the Analysis at 72 hr)	Analysis vs. Spin-Up at 72 hr
29° N			
T (K/10)	22/21/21	14/16/12	19/18/17
U (m s ⁻¹ /10)	48/50/53	43/43/40	20/24/27
V (m s ⁻¹ /10)	51/73/45	43/65/44	24/25/23
RH (%)	24/25/26	20/23/24	13/11/13
49° N			
T (K/10)	38/38/32	19/24/18	30/24/24
U (m s ⁻¹ /10)	72/53/86	49/42/63	42/31/42
V (m s ⁻¹ /10)	78/70/64	63/56/45	38/31/36
RH (%)	41/25/29	35/22/22	19/13/17
	Truth vs. Forecast at 120 hr (6-h GSM Forecast from the 42nd hr of the GDA)	Truth vs. Analysis at 120 hr (48th hr of the 48-hr GDA)	Analysis vs. Forecast at 120 hr
29° N			
T (K/10)	15/15/14	15/15/15	14/12/11
U (m s ⁻¹ /10)	62/54/51	62/52/45	30/32/32
V (m s ⁻¹ /10)	45/53/47	49/51/43	31/32/38
RH (%)	26/28/28	23/24/26	13/12. 13
49° N			
T (K/10)	12/12/17	16/10/15	14/11/13
U (m s ⁻¹ /10)	43/42/54	43/35/44	23/24/26
V (m s ⁻¹ /10)	45/35/54	37/33/49	24/19/27
RH (%)	21/24/26	20/22/27	12/13/14

A revealing analysis is obtained by subtracting the BASELINE OSSE from the CONTROL OSSE. Specifically, take the elements from the 16 rows and 3 columns of Tables 15 and 16 to form matrix [C15] and matrix [C16], respectively, and then calculate [C16]-[C15]. The results, which must be interpreted with care to ensure that only similar quantities are compared, are shown in Table 17.

Table 17 shows that the BASELINE and CONTROL spin-up forecast errors are identical, which must be the case because no assimilation occurred. The analysis errors (truth vs. the analysis made with the 72-hr GSM first guess forecast) are greater for the CONTROL than the BASELINE OSSE, as expected. The first guess was the same, so the differences are due only to the assimilation of simulated observations at the beginning of the GDA.

The differences between the analysis and the spin-up forecast of the CONTROL OSSE are also larger than those of the BASELINE OSSE, except occasionally for the meridional wind and relative humidity. The spin-up forecast is the same for BASELINE and CONTROL; thus, the introduction of errors into a set of perfect observations clearly has a measurable influence when the set of observations is analyzed.

The comparison between the BASELINE and CONTROL OSSEs is more interesting following the global data assimilations. The 6-hr GSM forecast errors (forecast vs. truth) from the 42nd hr of the GDA (that is, at the 120th hr of the three forecast/analysis cycles) are much larger in the CONTROL than the BASELINE OSSE, especially for wind components in lower latitudes. This result in general is expected because the sequential 6-hr forecasts in the BASELINE OSSE had the benefit of initial conditions that were closer to the truth values. Nevertheless, the difference between analysis error of zonal wind (nearly 2 m/s) and meridional wind (nearly 1.5 m/s) along the 29° latitude circle is relatively large.

The treatment of humidity by the CONTROL OSSE is difficult to understand. In some cases, at higher latitudes it appeared that assimilating observations with errors improved the forecast/analysis! Also, following the GDA, the RMS errors for the analysis and forecast are almost the same in the BASELINE as they are in the CONTROL. Based on this limited analysis, a reasonable conclusion is that in the GSM/ASAP system humidity observations have a relatively small impact on the resulting analyses, even if the observations are perfect.

Table 17. The Difference Between the BASELINE OSSEs, Which Assimilate "Perfect" Observations, and the CONTROL OSSEs, Which Assimilate Simulated Observations. The results are shown as the CONTROL OSSEs minus the BASELINE OSSEs; the format is the same as Tables 11 and 12.

Latitude/ Element	Forecast Error (Truth vs. 72-hr Spin-Up)	Analysis Error (Truth vs. the Analysis at 72 hr)	Analysis vs. Spin-Up
29° N			
T (K/10)	0/0/0	0/4/0	1/1/1
U (m s ⁻¹ /10)	0/0/0	0/3/0	2/2/2
V (m s ⁻¹ /10)	0/0/0	2/13/3	3/-1/0
RH (%)	0/0/0	0/2/0	1/0/-1
49° N			
T (K/10)	0/0/0	2/-1/1	1/0/0
U (m s ⁻¹ /10)	0/0/0	1/2/6	1/1/0
V (m s ⁻¹ /10)	0/0/0	6/4/4	-1-1/0
RH (%)	0/0/0	-1/1/1	2/-3/1
	Truth vs. Forecast at 120 hr (6-hr Forecast from GDA)	Truth vs. Analysis at 120 hr	Analysis vs. Forecast at 120 hr
29° N			
T (K/10)	4/2/2	5/4/5	7/4/4
U (m s ⁻¹ /10)	24/14/9	27/17/14	21/14/9
V (m s ⁻¹ /10)	10/15/9	15/15/2	18/7/11
RH (%)	7/7/3	4/4/4	5/2/1
49° N			
T (K/10)	1/1/4	6/0/6	6/1/3
U (m s ⁻¹ /10)	11/4/2	11/3/0	8/9/10
V (m s ⁻¹ /10)	7/2/3	7/2/3	8/6/6
RH (%)	-1/-2/2	2/-2/5	0/0/2

2.4.4 An Analysis of the TACOBS OSSE Global Data Assimilation Fields

The TACOBS scenario was developed by denying simulated data (the CONTROL case) from a selected portion of each of the two RAP regions (whose corners are 75°N 30°E, 45°N 30°E, 45°N 135°E and 75°N 135°E in the Eurasian region and 23°N 75°W, 23°N 120°W, 5°N 120°W, and 5°N 75°W in the central American region) as shown in Figs. 13 and 14, respectively. The GSM and the RWM were initialized with the 48th-hr of the appropriate GDAs and executed from all scenarios for 36 hr, producing forecasts at 6-hr intervals. The RWM, as usual, produced 12-, 18- (even though the forecast error correlation models were unsuccessful), 24- and 36-hr forecasts, the first guesses for RAP, for all scenarios.

The development and results of the global data assimilations of perfect (BASELINE case) and simulated (CONTROL case) observations have been described above. The simulated data-denied (TACOBS case) is also interesting. Visual comparisons of plots of the various meteorological fields resulting from the OSSEs revealed mostly trivial differences between the various scenarios of GDAs. Apparently, the effects of data denial in the eastern parts of the regions were felt downstream. Tables 18 and 19 show the rms differences, which provide useful objective comparisons of the GDAs.

A comparison of plots (in the RAP windows) of the meteorological fields resulting from a 48-hr GDA of TACOBS to those from a 48-hr GDA of CONTROL revealed differences between the two that were smaller than expected. The patterns of the meteorological fields in the eastern portion of the Eurasian region had much less detail in the TACOBS scenario, but most of the main features were recognizable. The fields were quite flat in the southern portion of the central American region in the BASELINE and CONTROL scenarios; so the lack of noticeable change in the TACOBS scenario perhaps should have been expected. Nevertheless, the rms differences between TACOBS and CONTROL illustrate that distinct changes occurred.

Table 18 shows the TACOBS scenario, whose GDA has characteristics similar to the BASELINE (Table 15) and CONTROL (Table 16) scenarios. Thus, the assimilation system showed little response to changes in errors assigned to observations and to changes in the number of observations being assimilated.

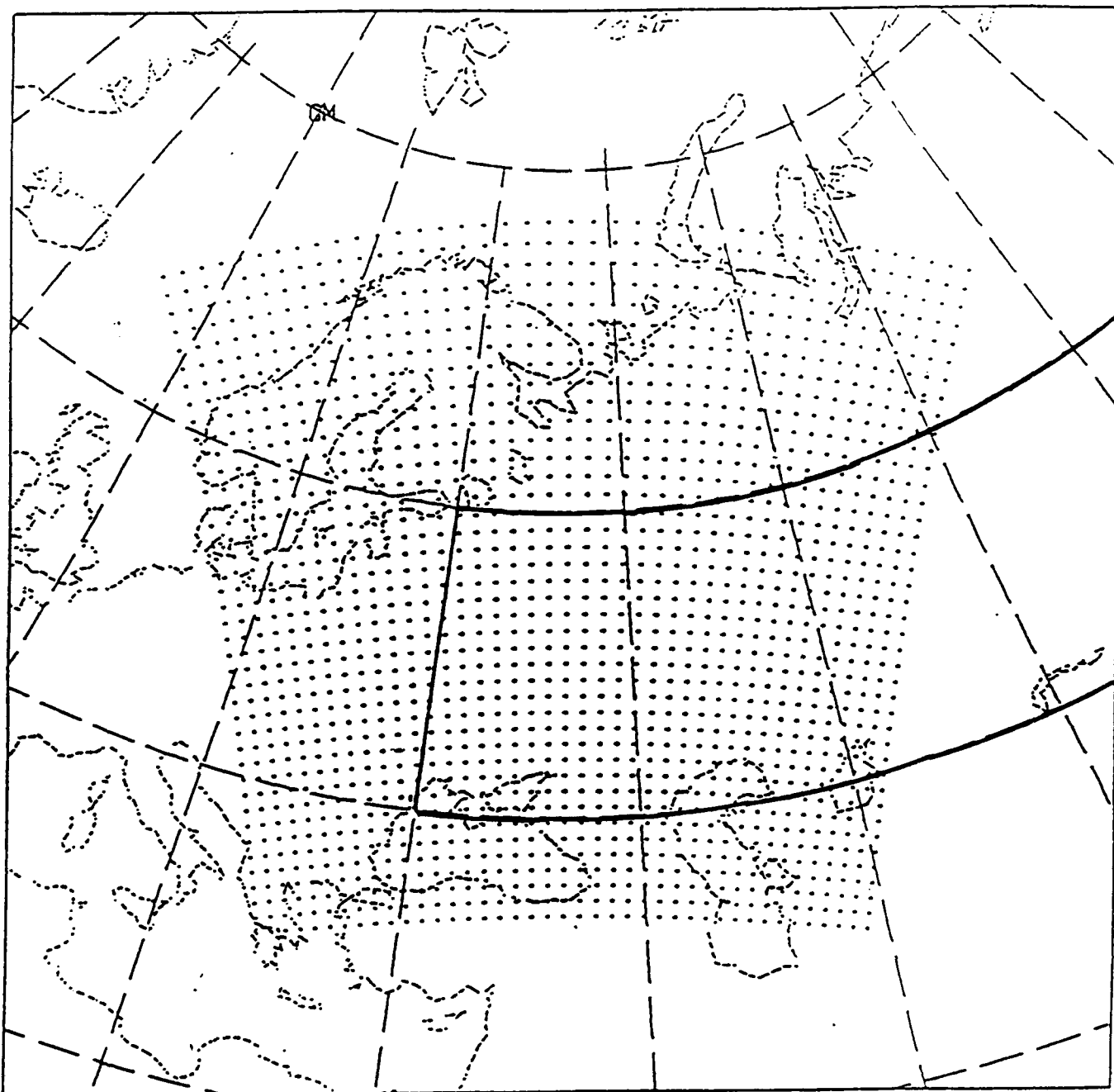


Figure 13. Same as Figure 1 with the addition of a heavy line surrounding the data-denied region of the TACOBS scenario in Eurasia.

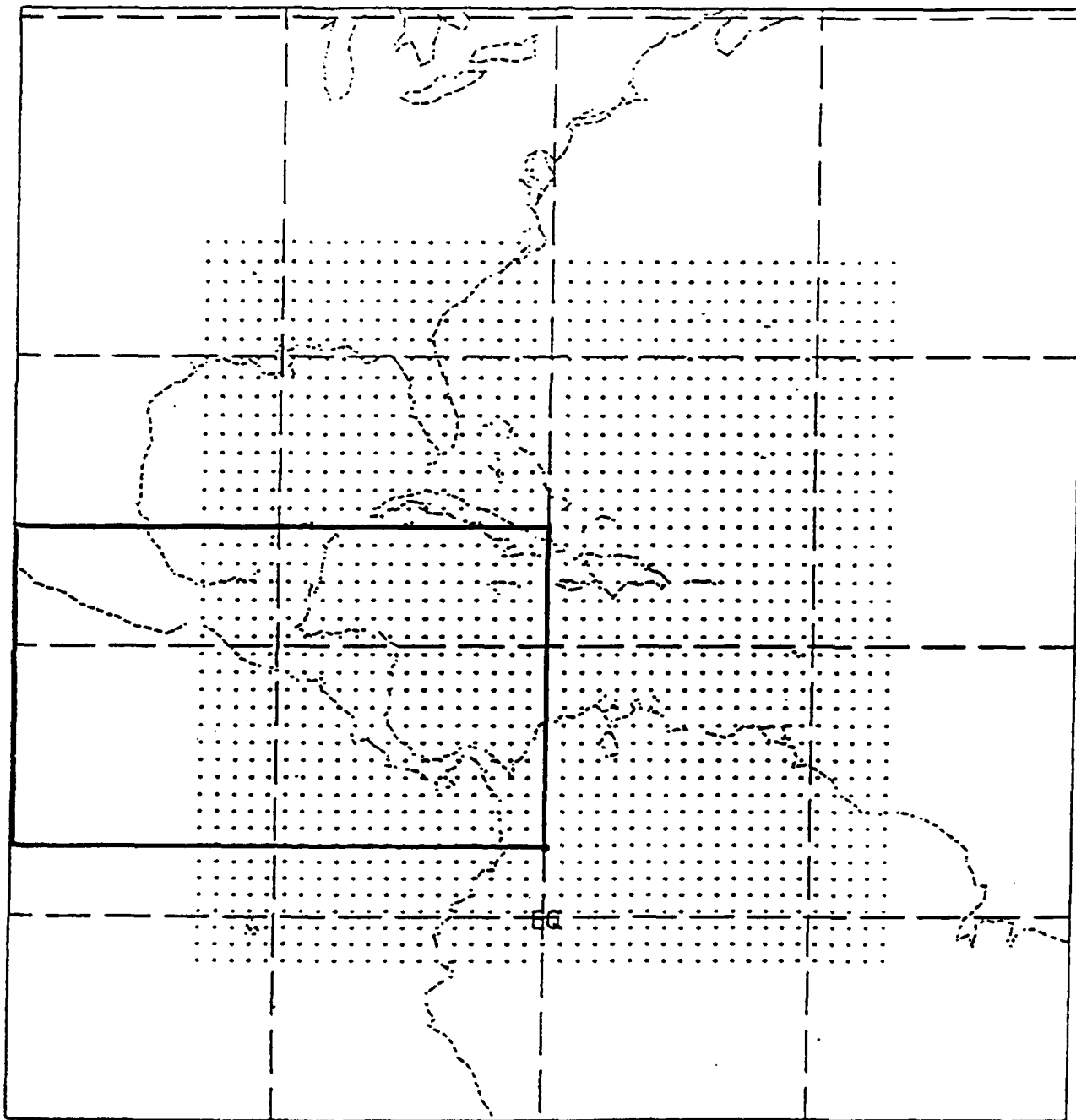


Figure 14. Same as Figure 2 with the addition of a heavy line surrounding the data-denied region of the TACOBS scenario in central America.

Table 18. The Root-Mean-Square Differences in the TACOBS OSSE Near 500 mb. Three OSSEs, Each Separated by a Slant, Begin at 00 UTC on 16 January 1979/ 21 January 1979/ 26 January 1979 (that is, T-106 hours 00/120/240) and Last for 120 hr. A 72-hr GSM forecast is followed by a 48-hr global data assimilation. The TACOBS scenario was denied all simulated observations (the CONTROL case) from two areas as shown in Figures 10 and 11.

Latitude/ Element	Forecast Error (Truth at 72 hr vs. 72-hr Spin-Up)	Analysis Error (Truth vs. the Analysis at 72 hr)	Analysis vs. Spin-Up at 72 hr
29°N			
T (K/10)	22/21/21	15/16/13	19/17/17
U (m s ⁻¹ /10)	48/50/53	43/44/41	20/24/27
V (m s ⁻¹ /10)	51/73/45	44/65/43	22/24/23
RH (%)	24/25/26	20/21/25	13/11/14
49°N			
T (K/10)	38/38/32	28/26/19	24/20/22
U (m s ⁻¹ /10)	72/53/86	52/42/63	43/38/38
V (m s ⁻¹ /10)	78/70/64	66/57/51	30/32/31
RH (%)	41/25/29	39/22/26	18/14/13
	Truth vs. Forecast at 120 hr (6-hr GSM Forecast from the 42nd hr of the GDA)	Truth vs. Analysis at 120 hr (48th hr of the 48-hr GDA)	Analysis vs. Forecast at 120 hr
29°N			
T (K/10)	14/14/16	14/15/17	14/12/13
U (m s ⁻¹ /10)	61/49/46	63/44/40	31/35/30
V (m s ⁻¹ /10)	56/56/50	56/54/47	35/34/38
RH (%)	22/24/29	21/21/28	13/10/12
49°N			
T (K/10)	26/20/21	26/18/20	23/16/15
U (m s ⁻¹ /10)	86/44/62	77/38/60	33/28/24
V (m s ⁻¹ /10)	82/38/56	69/40/53	33/19/28
RH (%)	27/26/29	26/27/27	15/16/17

Table 19. The Difference Between the TACOBS and CONTROL (TACOBS-CONTROL) OSSEs at Approximately 500 mb. The format is the same as Table 18.

Latitude/ Element	Forecast Error (Truth at 72 hr vs. 72-hr Spin-Up)	Analysis Error (Truth vs. the Analysis at 72 hr)	Analysis vs. Spin-Up at 72 hr
29°N			
T (K/10)	0/0/0	1/0/1	1/-1/0
U (m s ⁻¹ /10)	0/0/0	0/1/1	0/0/0
V (m s ⁻¹ /10)	0/0/0	1/0/-1	-2/-1/0
RH (%)	0/0/0	0/-2/1	0/0/1
49°N			
T (K/10)	0/0/0	9/2/1	-6/-4/-2
U (m s ⁻¹ /10)	0/0/0	3/0/0	1/7/-4
V (m s ⁻¹ /10)	0/0/0	0/1/6	-8/1/-5
RH (%)	0/0/0	4/0/4	-1/1/-4
	Truth vs. Forecast at 120 hr (6-h GSM Forecast from the 42nd hr of the GDA)	Truth vs. Analysis at 120 hr (48th hr of the 48-hr GDA)	Analysis vs. Forecast at 120 hr
29°N			
T (K/10)	-1/-1/2	-1/0/2	0/0/2
U (m s ⁻¹ /10)	-1/-5/-5	1/-8/-5	1/3/-2
V (m s ⁻¹ /10)	11/3/3	7/3/4	4/2/0
RH (%)	-4/-4/1	2/-3/2	0/-2/-1
49°N			
T (K/10)	12/8/4	10/8/5	9/5/2
U (m s ⁻¹ /10)	44/2/7	34/3/26	10/4/1
V (m s ⁻¹ /10)	37/3/2	22/7/4	9/0/1
RH (%)	6/2/3	6/5/0	3/3/3

The best way to examine the changes between scenarios, however, is to form a table of the differences between two scenarios (similar to Table 15) by subtracting CONTROL from TACOBS. Table 19 shows the differences between the TACOBS and the CONTROL scenarios. Several interesting features are immediately apparent. The experiment beginning on 16 January 1979 was most strongly affected by the data denial, especially the northern region (i.e., along the 49°N latitude circle) where very large differences in the wind field occurred. For all three experiments there, the TACOBS errors were larger than the CONTROL errors, as expected, because fewer observations were available for assimilation. Relative humidity was least affected, reconfirming the hypothesis that the forecast/GDA cycle is weakly sensitive to changes in humidity observations.

Recall that the GDA numerical verification took place around entire latitude circles; therefore, the substantial effect of withholding observations from an analysis was readily detected downstream. On the other hand, the visual comparisons provided from the view of the RAP windows could not have been expected to be the complete picture. Nevertheless, the visual comparisons were representative of what an observer in the RAP region would have experienced, that is, the relatively small effect of data denial.

The southern region had much smaller differences between the two scenarios than the northern region. The southern region also had the strange result of the TACOBS scenario having typically smaller errors from the GDA in all variables except the meridional wind (recall that the wind fields had very large errors). The removal of the observations from the west-central portion of the central American region apparently slightly improved the results of the GDA.

It is interesting to note from a glance at Tables 17 and 19 that in the comparatively data dense northern region the effect of observation errors at the 48th hr of the GDA is much less than the effect of missing observations. In the comparatively data sparse southern region, however, the effect of observation errors was larger than missing observations. This may be due to the relatively small number of observations missing from the southern region.

2.4.5 The RWM Forecasts from the OSSEs

The 48th-hr of the GDA of each of the three observation scenarios was used to initialize the GSM (which generated forecasts at 6-hr intervals for the boundary conditions for the RWM). The RWM was run for 36 hr to produce the 12-, 18-, 24- and 36-hr forecasts, first guesses for RAP. The results from a visual comparison of truth to plots of 500-mb RWM 36-hr forecasts initialized from the BASELINE, CONTROL, and TACOBS scenarios are discussed below:

Central American Region

Verification on 22 Jan 79, 12 UTC

The major feature of the geopotential field was a 521-dm low east of Cape Cod and south of Yarmouth, Nova Scotia. Strong ridging occurred over the western Great Lakes and the northeastern corner of the region. All three scenarios had similar forecasts of a low of 532 dm. The low in both the BASELINE and CONTROL scenario was positioned close to truth, but the low in the TACOBS scenario was at least 200 km too far south.

The jet streak (41.6 m/s) in the zonal wind was located several hundred kilometers south of the 500-mb low, about 1000 km due east of Cape Hatteras. The BASELINE and CONTROL scenarios were similarly placed but 2 and 4 m/s, respectively, too strong. The TACOBS jetstreak was only 2 m/s too strong but was about 200 km too far south.

The meridional wind field had a northerly flow of 36 m/s over western NY and a southerly flow of 32 m/s centered over 38°N and 60°W. Both the BASELINE and CONTROL meridional flows were properly positioned but 7 m/s too weak. The TACOBS northerly jet was only 1 m/s too weak and shifted slightly too far southeastward; the southerly jet was well placed but was 8 m/s too weak.

The basic relative humidity patterns of the BASELINE and CONTROL scenarios were similar to truth, except for too much moisture near the Equator just east of Equador and associated with the southerly jet in the western Atlantic. TACOBS was much too moist in the Gulf of Mexico and the western Caribbean Sea.

Eurasian Region

Verification on 22 Jan 79, 12 UTC

The main geopotential features were a 544-dm low just east of the Sea of Azov, a short wave over Finland, and a low east of the region in Siberia. The three forecast scenarios were similar, especially in the western portion of the region.

The zonal wind forecasts were good along 50-60°N but verified poorly to the north and south in all scenarios. As expected, the BASELINE forecast was best and the TACOBS forecast was worst; nevertheless, all forecasts were similar.

The meridional wind had maxima from the north over northern Scandinavia (-18 m/s) and the Central Urals (-26 m/s) and secondary maxima surrounding the Black Sea. The forecast scenarios were similar: good in the vicinity of the Black Sea, about 6 m/s too weak over the Urals, and poor over Scandinavia. The TACOBS forecast over the Urals was the best.

The relative humidity forecasts replicated the main features of truth. However, in general they were too moist.

Central American Region

Verification on 27 Jan 79, 12 UTC

The major feature of the geopotential field was a 521-dm low east of Cape Cod and south of Yarmouth, Nova Scotia. This low was almost of the exact same depth and location as 5 days earlier. None of the forecasts, however, picked up this system. Weak ridging occurred over the western Great Lakes, but all the forecasts had a basic zonal flow.

The core of the zonal jet (48 m/s) was located off the coast of Virginia and south of the low center. The low caused a strong easterly flow north of 40°N. The BASELINE forecast core was displaced several hundred kilometers to the west-northwest, the CONTROL forecast core was displaced about 300 km to the northwest, and the TACOBS forecast was only about 200 km too far north. It is interesting that the TACOBS scenario (with data denied in the southwestern portion of the region) was the best forecast.

The meridional wind field had a strong southerly jet (22 m/s) south of Newfoundland and a northerly jet (-19 m/s) from New York State to Virginia. All forecasts failed over the northeastern portion of the region.

The relative humidity forecasts over the northeast were too dry.

Eurasian Region

Verification on 27 Jan 79, 12 UTC

The main features of the geopotential fields were a 504-dm low over the English Channel, ridging from the Black Sea northward to Scandinavia, and a weak trough over the Caspian Sea. All forecasts were similarly good, although only the TACOBS forecast correctly did not cut off a low over the Caspian Sea.

Strong zonal flow extended north of the 60°N latitude circle with a core of 26 m/s over the White Sea. Weak easterlies occurred over and east of the Caspian Sea. All forecasts were similar to truth.

A strong (34 m/s) meridional flow was centered over Denmark, extending southeastward to Turkey and north of the Arctic Circle. A weaker (-15 m/s) northerly flow was over much of Siberia. The forecasts were best in the western part of the region; but even the details of the eastern portion of the region were captured reasonably well. The TACOBS forecast, as might be expected, had problems in the east, where data had been denied.

The relative humidity forecasts represented the major patterns. The TACOBS forecast, however, was not good in the east.

Central American Region
Verification on 1 Feb 79, 12 UTC

The 500-mb pattern had relaxed to a broad trough over the east and a sharp ridge in the northeast portion of the region (south of Newfoundland). The forecasts of this relatively simple pattern were very good.

A strong zonal flow with a jet core of 49 m/s dominated the central United States and turned to the northeast as it reached the Atlantic. The BASELINE forecast, however, had a westerly jet of 55 m/s from the Great Lakes to east of Cape Cod. The CONTROL forecast was similar, but the TACOBS forecast was weaker by 3 m/s.

The meridional jet was 30 m/s east of Cape Cod and extended north of Nova Scotia. All forecast scenarios failed to capture this feature.

The relative humidity was high over the western Gulf of Mexico. Both the CONTROL and TACOBS forecast scenarios had this feature, but not the BASELINE. The dry area south of Cape Cod and the moist area to the northeast were forecasted in all scenarios.

Eurasian Region
Verification on 1 Feb 79, 12 UTC

The major features in the geopotential field were a 515-dm low over Denmark, a 548-dm ridge centered over the Barents Sea, and 516-dm low over north-central Siberia. All three forecasts cut off a 554-dm (within 10 m) high northwest of the Barents Sea and displaced the weakened Siberian low far to the southeast.

The zonal wind field was relatively weak. All three forecasts were similar and verified well.

The more interesting meridional wind field has a strong southerly flow (20 m/s) over Europe and Scandinavia, and strong northerly flow (30 m/s) over the eastern Barents Sea as well as a weaker northerly flow (19 m/s) over central Eurasia. The CONTROL and BASELINE forecasts were good except

for missing the forecast of the jet over the Kara Sea. The TACOBS forecast missed all features over eastern Eurasia.

The relative humidity pattern was especially complex. However, the forecasts verified unexpectedly well, even the TACOBS.

2.4.6 Results of RAP Analyses from the OSSEs

This subsection consists of Tables 20-34 that show some of the results of the OSSEs at 500 mb (similar tables at 850 mb and 300 mb are provided in Appendix B.) The following RWM forecast errors (defined as the difference between the gridded forecast and T-106 truth field) and RAP analyses errors (defined as the difference between the gridded analysis and T-106 truth field) were produced: height of mandatory pressure levels (1000 - 50 mb) and the associated temperature, dewpoint, zonal wind component (u), and meridional wind component (v). The errors are grouped within the 2 regions, 12 levels, and 5 variables.

Each GDA, forecast, and RAP analysis had access only to the set of observations that was consistent with the specific OSSE. In the TACOBS OSSE, for example, the TACOBS data set was used throughout the 48-hr GDA, which initialized the GSM/RWM forecasts, and only the TACOBS observations were available to RAP.

Recall that the central American region was treated twice, once with satellite data included and then with satellite data excluded from the set of available observations; the Eurasian region was treated only once because satellite data had no influence there. Thus, there are possibly 180 (3 regions x 12 levels x 5 variables) unique tables to present, from which a representative sample of results of the OSSEs is included here. Many of the 180 tables show nearly redundant information, especially at adjacent mandatory levels; therefore, only low (850 mb), middle (500 mb), and high (300 mb) levels are included.

Each table allows a rapid comparison of the three forecast models (12, 24, and 36 hr), the three scenarios (BASELINE, CONTROL, and TACOBS), and the three analysis periods, each separated by 5 days. The tables show errors in a region for a given meteorological element on a mandatory level, and thereby provide confidence limits of RAP analyses. An entry of NA means that the data are not available (the project ended before some production jobs could be resubmitted).

Table 20. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included over Eurasia (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	6.835	15.47	NA	NA	6.630	29.53
24 hr BASELINE	6.892	21.91	NA	NA	7.833	35.89
36 hr BASELINE	7.012	35.61	NA	NA	8.558	39.91
12 hr CONTROL	10.58	16.80	9.229	58.42	10.09	29.77
24 hr CONTROL	10.83	23.10	10.17	50.39	12.30	35.82
36 hr CONTROL	12.01	38.35	11.88	42.56	11.95	42.45
12 hr TACOBS	21.20	22.16	17.41	62.19	16.55	26.37
24 hr TACOBS	25.66	32.18	16.83	52.32	19.99	29.99
36 hr TACOBS	28.73	47.31	20.69	42.19	23.62	34.34

In general and as expected, the analysis errors of the height of the 500-mb surface increase with the length of the first guess forecast for all scenarios and observation sets. As measured by the ratio of the analysis errors to forecasts errors, however, the reliability of the RAP analyses is not always a strong function of the first guess, especially for the 31 January CONTROL case. An older first guess is still effective, but not as good as a more recent forecast.

Table 21. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess over Eurasia Forecasts on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included over Eurasia (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.3585	1.207	NA	NA	0.4048	0.9164
24 hr BASELINE	0.4275	1.625	NA	NA	0.4697	1.238
36 hr BASELINE	0.4355	2.238	NA	NA	0.4978	1.776
12 hr CONTROL	1.430	1.221	1.292	1.271	1.092	.9640
24 hr CONTROL	1.521	1.593	1.174	1.860	1.368	1.405
36 hr CONTROL	1.266	2.190	1.299	2.376	1.473	2.089
12 hr TACOBS	2.170	2.250	1.560	1.762	2.034	2.114
24 hr TACOBS	1.975	2.391	2.108	2.239	2.431	2.566
36 hr TACOBS	1.675	2.844	2.254	2.718	2.386	2.900

The BASELINE temperature analysis errors also increase with the length of the first guess forecast. The ratio of analysis errors to forecast errors is, however, larger for temperature than for heights. The CONTROL temperature analysis made from the 12-hr forecast is inferior to the first guess for all analyses, the analysis made from the 24-hr forecasts is slightly better than the first guess, and the analysis made from the 36-hr forecast is clearly superior to the first guess. The TACOBS temperature analysis is similar to the 12- and 24-hr first guesses, and better than the 36-hr forecast. The simulated analyses appear to have used older forecasts very effectively.

Table 22. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.669	3.613	NA	NA	2.674	3.513
24 hr BASELINE	2.905	4.169	NA	NA	3.204	4.223
36 hr BASELINE	3.178	5.327	NA	NA	3.539	5.750
12 hr CONTROL	4.307	3.867	2.315	3.212	4.458	4.235
24 hr CONTROL	3.407	4.012	3.362	4.346	3.958	5.052
36 hr simulated	3.456	5.211	4.513	6.196	4.898	6.006
12 hr TACOBS	4.668	5.281	3.647	3.874	5.135	5.390
24 hr TACOBS	4.945	5.533	4.426	4.760	4.934	5.800
36 hr TACOBS	4.858	6.364	5.762	5.949	6.529	6.457

The dewpoint errors are similar to the temperature errors except for the BASELINE scenario, where the analysis errors are much larger when compared to the first guess forecast. Experiments with real data suggest that the dewpoint forecast error correlation needed to be improved, and the BASELINE experiment confirms the need. Nevertheless, even with simulated observations and data denial, the models provide analyses that are better than the first guess forecasts.

Table 23. Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.408	2.447	NA	NA	1.048	2.746
24 hr BASELINE	1.477	3.463	NA	NA	1.733	3.531
36 hr BASELINE	1.203	4.670	NA	NA	1.937	4.380
12 hr CONTROL	3.593	2.916	3.874	2.315	4.033	3.285
24 hr CONTROL	3.869	3.619	3.937	2.842	4.251	4.110
36 hr CONTROL	3.771	4.803	3.653	3.596	3.892	5.099
12 hr TACOBS	3.998	3.235	4.296	2.920	4.054	4.341
24 hr TACOBS	4.547	4.591	3.774	3.239	5.124	5.226
36 hr TACOBS	4.995	6.029	3.994	3.719	5.010	5.883

The ratio of analysis errors of the zonal wind component to its first guess in the BASELINE scenario is comparable to similar ratios of height and temperature, suggesting that the forecasting error models are of similar quality. The reliability of the RAP analysis is also a similar function of the length of the first guess forecast. Only the analyses that used a 36-hr first guess are better than the first guess forecasts. Note that the CONTROL analysis errors are 250% larger than the BASELINE errors, while the first guess forecast errors are only slightly larger by comparison.

Table 24. Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.281	2.498	NA	NA	1.091	2.274
24 hr BASELINE	1.514	3.416	NA	NA	1.812	3.098
36 hr BASELINE	1.275	4.615	NA	NA	1.797	4.467
12 hr CONTROL	4.178	3.086	3.389	2.305	4.377	2.493
24 hr CONTROL	3.875	3.912	4.194	3.122	4.262	3.782
36 hr CONTROL	3.613	5.130	3.387	3.668	3.738	4.995
12 hr TACOBS	5.325	4.654	3.872	2.950	5.319	3.413
24 hr TACOBS	5.601	5.656	4.004	3.424	4.374	3.823
36 hr TACOBS	5.147	6.768	3.194	3.658	4.402	4.408

The analysis errors of the meridional wind component are similar to the analysis errors of the zonal wind component. The errors are comparatively (ratio of analysis errors to forecast error) much larger than the errors of other analyzed variables. Since the quality of the CONTROL first guess forecasts was not notably different from the BASELINE forecasts, the observations in the CONTROL scenario appear to have large errors.

Table 25. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	10.72	25.94	9.135	30.67	8.400	24.47
24 hr BASELINE	13.88	31.32	12.03	34.32	7.419	20.50
36 hr BASELINE	15.33	37.30	14.56	41.14	8.175	16.95
12 hr CONTROL	13.94	26.28	12.69	32.38	14.67	23.31
24 hr CONTROL	14.41	28.75	13.31	34.77	11.83	22.76
36 hr CONTROL	16.76	36.54	16.63	41.80	14.05	17.01
12 hr TACOBS	14.89	26.02	17.25	31.69	16.96	26.06
24 hr TACOBS	14.75	28.10	15.53	33.72	13.32	24.52
36 hr TACOBS	18.69	35.90	17.38	40.73	18.59	17.39

The BASELINE and CONTROL height analysis errors are somewhat larger in the central American region than in the Eurasian region, even though the forecast errors are similar. The ratio of the BASELINE analysis errors to the first guess forecast errors is also larger, especially for the 36-hr forecast. Nevertheless, the typical height errors are small (less than 20 m).

Table 26. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.8184	1.600	0.7361	1.475	0.8173	0.9536
24 hr BASELINE	1.148	1.859	0.9187	1.828	0.9025	1.101
36 hr BASELINE	1.347	2.043	0.9137	1.949	0.7624	1.259
12 hr CONTROL	1.304	1.716	1.112	1.511	1.274	1.130
24 hr CONTROL	1.210	1.844	1.059	1.958	1.330	1.223
36 hr CONTROL	1.404	2.109	1.199	2.083	1.306	1.278
12 hr TACOBS	1.232	1.665	1.335	1.631	1.221	1.039
24 hr TACOBS	1.264	1.814	1.515	1.978	1.346	1.144
36 hr TACOBS	1.432	2.028	1.377	2.197	1.382	1.254

The BASELINE temperature analysis errors are larger in central America than in Eurasia; the ratio of analysis error to forecast error is nearly twice as large. The CONTROL analysis errors in each region are similar, but the TACOBS analysis errors are smaller. In the 31 January experiment, the BASELINE analysis errors were only a slight improvement of the first guess, and the CONTROL and TACOBS analyses were worse than the first guesses.

Table 27. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	7.923	10.27	6.868	9.323	6.316	7.719
24 hr BASELINE	7.523	10.31	6.428	9.219	7.308	8.214
36 hr BASELINE	8.416	10.37	6.531	9.136	5.208	7.677
12 hr CONTROL	9.956	10.31	8.979	11.02	8.764	8.728
24 hr CONTROL	9.990	10.04	9.811	10.70	9.461	8.984
36 hr CONTROL	11.86	9.918	10.13	10.43	8.033	8.094
12 hr TACOBS	10.33	10.98	9.804	11.12	8.834	9.147
24 hr TACOBS	10.63	10.78	11.57	10.78	9.620	9.346
36 hr TACOBS	11.95	10.40	12.74	10.74	7.405	8.561

The dewpoint analysis from all observation scenarios are similar in both regions when compared their first guess forecasts; however, the magnitude of the errors is much larger in central America, almost by a factor of 3. The dewpoint forecasts in central America are poor; they dominate the analyses, which are even worse than the forecast in the TACOBS scenario.

Table 28. Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.437	2.864	2.782	2.883	2.181	2.517
24 hr BASELINE	2.708	2.734	2.855	3.697	2.806	3.291
36 hr BASELINE	3.619	3.713	3.249	5.077	2.557	4.813
12 hr CONTROL	4.658	3.694	4.682	3.677	5.159	3.063
24 hr CONTROL	4.376	3.648	5.243	4.289	5.974	3.356
36 hr CONTROL	5.786	4.677	4.343	4.911	4.199	4.761
12 hr TACOBS	4.523	3.373	5.615	3.891	5.169	3.406
24 hr TACOBS	5.037	3.349	5.406	4.110	6.866	4.012
36 hr TACOBS	5.712	4.065	5.024	4.929	4.315	5.487

The zonal wind analysis errors in central America are larger than in Eurasia, especially the BASELINE analysis whose errors are twice as large. The first guess forecasts are of comparable quality; therefore, the forecast error correlation models for central America need to be improved, especially the 12-hr model. (For all scenarios the 12-hr model produced the best analysis in only two of nine cases.) The CONTROL and TACOBS scenarios, however, have analyses that are consistently worse than their first guess forecasts.

Table 29. Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.651	2.940	1.933	3.463	2.118	2.575
24 hr BASELINE	2.373	3.151	3.362	5.369	2.072	2.781
36 hr BASELINE	2.582	4.393	2.226	5.131	2.094	3.423
12 hr CONTROL	4.435	4.078	4.294	3.863	4.910	3.293
24 hr CONTROL	5.430	3.409	4.978	5.348	4.934	3.245
36 hr CONTROL	5.439	4.308	5.819	4.938	4.758	3.926
12 hr TACOBS	4.446	3.649	5.328	4.293	5.130	2.911
24 hr TACOBS	5.581	3.033	4.585	5.402	5.862	3.328
36 hr TACOBS	6.071	3.913	5.605	5.208	5.403	3.931

The BASELINE meridional wind analysis errors are slightly larger in central America than in Eurasia, but the forecast errors are similar in both regions. Based upon the ratio of analysis errors to forecast errors in central America, the meridional wind forecast error correlation model appears to be better than the zonal model. The CONTROL and TACOBS analysis errors in central America are similar to those in Eurasia, although the effects of data denial are more apparent in Eurasia. Once again, the analyses from simulated wind observations are much worse than the first guess forecasts.

Table 30. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	8.699	25.94	9.130	30.67	8.483	24.47
24 hr BASELINE	13.92	31.32	12.01	34.32	7.419	20.50
36 hr BASELINE	13.86	37.30	14.55	41.14	8.227	16.95
12 hr CONTROL	12.43	26.28	12.68	32.38	14.84	23.31
24 hr CONTROL	14.42	28.75	13.25	34.77	11.83	22.76
36 hr CONTROL	15.08	36.54	16.59	41.80	14.11	17.01
12 hr TACOBS	14.65	26.02	17.59	31.69	17.39	26.06
24 hr TACOBS	14.99	28.10	15.56	33.72	13.32	24.52
36 hr TACOBS	17.09	35.90	16.77	40.73	19.89	17.39

As can be readily seen by comparing Tables 25-29 with Tables 30-34, respectively, excluding satellite data had hardly any effect on most analyses. The only notable effects occurred during the 21 January analysis of height, when the height errors were reduced by 2 m. Clearly, a better approach to incorporating satellite data into analyses is required.

Table 31. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.8557	1.600	0.7338	1.475	0.8077	0.9536
24 hr BASELINE	1.153	1.859	0.9099	1.828	0.9025	1.101
36 hr BASELINE	1.305	2.043	0.9152	1.949	0.7598	1.259
12 hr CONTROL	1.363	1.716	1.111	1.511	1.270	1.130
24 hr CONTROL	1.216	1.844	1.060	1.958	1.330	1.223
36 hr CONTROL	1.327	2.109	1.202	2.083	1.308	1.276
12 hr TACOBS	1.271	1.665	1.334	1.631	1.208	1.039
24 hr TACOBS	1.298	1.814	1.386	1.978	1.346	1.144
36 hr TACOBS	1.508	2.028	1.409	2.197	1.412	1.254

Table 32. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	7.923	10.27	6.868	9.323	6.316	7.719
24 hr BASELINE	7.523	10.31	6.428	9.219	7.308	8.214
36 hr BASELINE	8.294	10.37	6.531	9.136	5.208	7.677
12 hr CONTROL	9.956	10.31	8.979	11.02	8.764	8.728
24 hr CONTROL	9.990	10.04	9.811	10.70	9.461	8.984
36 hr CONTROL	11.81	9.918	10.13	10.43	8.033	8.094
12 hr TACOBS	10.33	10.98	9.804	11.12	8.834	9.147
24 hr TACOBS	10.63	10.78	11.57	10.78	9.620	9.346
36 hr TACOBS	11.68	10.40	12.65	10.74	7.186	8.561

Table 33. Root-Mean-Square U-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central America Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Excluded (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.437	2.864	2.782	2.883	2.177	2.517
24 hr BASELINE	2.708	2.734	2.855	3.697	2.806	3.291
36 hr BASELINE	3.570	3.713	3.249	5.077	2.557	4.813
12 hr CONTROL	4.574	3.694	4.682	3.677	5.172	3.063
24 hr CONTROL	4.376	3.648	5.243	4.289	5.974	3.356
36 CONTROL	5.746	4.677	4.343	4.911	4.199	4.761
12 hr TACOBS	4.482	3.373	5.614	3.891	5.149	3.406
24 hr TACOBS	5.037	3.349	5.406	4.110	6.866	4.012
36 hr TACOBS	5.700	4.065	5.016	4.929	4.434	5.487

Table 34. Root-Mean-Square V-wind Errors (m/s) of RAP Analyses and First Guess Forecasts over the Central American Region on the 500-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Not Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.678	2.940	1.934	3.463	2.119	2.575
24 hr BASELINE	2.373	3.151	3.362	5.369	2.072	2.781
36 hr BASELINE	2.546	4.393	2.226	5.131	2.094	3.423
12 hr CONTROL	4.389	4.078	4.295	3.863	4.909	3.293
24 hr CONTROL	5.430	3.409	4.978	5.348	4.934	3.245
36 hr CONTROL	5.431	4.308	5.819	4.938	4.758	3.926
12 hr TACOBS	4.360	3.649	5.314	4.292	5.127	2.911
24 hr TACOBS	5.581	3.033	4.585	5.402	5.862	3.328
36 hr TACOBS	6.042	3.913	5.730	5.208	5.541	3.031

2.5 STC TASK 5

The purpose of this task is to perform a survey of the literature on short-term forecasting and recommend a candidate technique to extend the RAP analysis to a 12-hr forecast. The survey was limited to literature from the past 10 years, the period following the Symposium on Nowcasting held in Hamburg, Germany in August 1981. The symposium formalized the concept of nowcasting: the diagnosis and depiction of the detailed distribution of weather and localized weather changes with a lead time of up to 12 hr.

In general, nowcasting attempts to assimilate all meteorological data, especially unconventional data (geostationary satellite imagery and atmospheric sounding, Doppler radar, weather radar, atmospheric profiler, etc.). It then forecasts the movement and development of weather systems.

2.5.1 Background

First, consider the obvious candidate, persistence of the analysis. That is, assume the fields present at time zero will remain unchanged for 12 hr. Further consideration suggests that extrapolating the fields with a specified advection velocity would be better than persistence.

Even for the relatively large RAP region, which is approximately 2000 km across, a weather system just outside its boundary at time zero could dominate its weather in 12 hr. An advection speed of only 10 m/s, for example, would allow a system to move 430 km into the region. Obviously, a greater advection speed would cause more of the region to be affected. If the forecast is required only at points far from the boundaries being affected by advection, extrapolation of the system would be a reasonable solution to the short-term forecasting problem. In general, however, better solutions are needed.

The RAP proposal offered a possible solution during a discussion of how the RWM forecasts, already available as the first guesses for RAP, can be modified to "extend" the analysis for 12 hr. After the RAP analysis has been completed at time zero, the difference field between the first guess and the analysis can be readily obtained. These time zero difference fields, which should be closely related to the differences that will occur 12 hr later between the forecast and the actual weather analysis, may need to be adjusted for the 12-hr time change (as discussed below). The RWM forecast, assumed to be available 12 hr beyond the analysis time, and the time-zero difference, adjusted for the 12-hr time change if necessary, can be added together. The result in effect is an extension of the RAP analysis, obtained by relatively simple means, of all variables at all grid points.

2.5.2 Extension of Solution Discussed in the Proposal

The simplest adjustment of the time-zero differences would, of course, be no adjustment. Thus, the differences between the first guess and the gridded analysis at time zero would be added directly to the forecast verifying at time zero plus 12 hr. Presumably, the same forecast procedure used to supply

the first guess at time zero could also supply a forecast for the subsequent 12 hr (and 24 hr for that matter), especially if the forecast procedure involved a numerical weather prediction (NWP) model.⁴

Time-zero differences in the analysis window will probably closely resemble the time-zero differences at the grid points 12 and even 24 hr later. If the NWP forecast at time zero is somewhat out of phase with its verification (the time-zero analysis), the 12- and 24-hr forecasts, which are based on the same initialization of the time-zero forecast, are likely to be similarly out of phase with their corresponding verifications. Furthermore, the same likelihood obtains with respect to mispredictions of the intensity of a weather system. Therefore, with the changes due to advection and development reasonably accounted for by the forecast model, the quasistatic first guess differences from time zero added to the 12- and 24-hr forecasts should yield reasonable 12- and 24-hr forecasts. These forecasts (or pseudoanalyses) can be derived with virtually no effort beyond that required for the time-zero analysis.

Improved pseudoanalyses may be obtained with some additional effort by determining the relationship between first guess analyses separated by 12 and 24 hr. Even though only modest changes in the first guess fields are expected over 12 and 24 hr, some of the changes will be systematic and can be taken into consideration. For example, time-zero differences could be adjusted based on knowledge of time-series behavior, which can be calculated from a sufficiently lengthy series of first guess fields. Separate studies would be required for each analysis variable, standard level, time of day, season, and geographic location. Once knowledge of the nature of the short-term changes in first guess forecast differences is available, the time-zero differences can be modified to yield improved pseudoanalyses.

2.5.3 Literature Survey

In *Nowcasting*, Beran and MacDonald (1982) point out the characteristics of very short-range forecasts (0-12 hr), also called nowcasts. These forecasts are for a regional area where spacing is measured in terms of tens of kilometers (mesoscale). Typical nowcasting methodologies include extrapolation by NWP models or statistical techniques. Beran and MacDonald noted that the capabilities of the NWP models, which have been remarkably successful at forecasting for synoptic scales, have been

⁴Without regard to the nature of the forecast procedure, however, it is assumed to provide reliable forecasts; otherwise, it would not have been chosen to supply the time-zero first guess. This assumption eliminates from consideration such quasistatic first guess alternatives as climatology. For purposes here the first guess is assumed to be supplied by a regional NWP model.

transferred only in special cases, such as hurricanes, to the mesoscale. This is generally true today, 10 years later.

Applying mesoscale modeling to forecasting synoptically induced mesoscale systems requires initialization with conventional observations as well as the details of precipitation (radar echoes) and satellite imagery (clouds and moisture). This is a difficult task that often yields results little better than extrapolation. In addition, Pielke (1981) noted that in the mesoscale many weather systems are terrain induced. Thus, a mesoscale NWP model should account for topography.

Nevertheless, some NWP models have had limited success (for example, Thompson, 1988; Warner and Seaman, 1990; and Mills and Seaman, 1990). Petersen and Homan (1989) described a simple, isentropic NWP model for making short-term forecasts. The purpose of this model is to extend the usefulness of the details present in the initial observations out to 12 hr. The Petersen and Homan model runs on a small computer and provides estimates of the changes of atmospheric temperature, moisture, and wind field. It was especially useful at advecting Visible and Infrared Spin-Scan Radiometer (VISSR) Atmospheric Sounder (VAS) high-resolution moisture imagery.

Browning (1982) believed much can be gained from a conceptually simple approach involving the detailed description of the current weather pattern and extrapolating it for several hours ahead. On the other hand he warned that forecasts based upon simple linear extrapolation become increasingly unreliable as the lead time extends beyond a few hr because of development and decay.

Mc Arthur, Davis, and Reynolds (1987) pointed out the advantages of pattern recognition in short-range forecasting. They constructed a knowledge-based system to guide the development of objective pattern recognition algorithms for use with meteorological data. Their system relies on expert-level nowcasters who depend heavily on the use of weather scenarios that characterize generic classes of weather patterns and their development over time. The main advantage is that the system allows for rapid focusing on the most useful data and ignoring the rest. The system is useful for forecasting specific meteorological events (such as thunderstorms) at a location. The authors believe that this concept in principle could be extended to year-round weather phenomena for a region.

Wilks (1991) noted that "knowledge of today's weather typically provides some information regarding tomorrow's weather, even absent a forecast for tomorrow," another way of saying that there is a positive serial correlation between forecasts of weather events. He successfully took this approach to forecasting a situation (precipitation), where the autocorrelation of the meteorological process is strong. This approach does not directly assist extending a RAP analysis but shows that using serial correlations could, as suggested in the background discussion (Section 2.5.1).

Van den Dool (1989) proposed using a limited-area analogue forecasting (AF) model for a 12-hr forecast at a point and then applying the AF procedure to a large number of adjacent grid points. He considered a 12-hr forecast of the 500-mb height at a grid point by drawing a circle (of radius $r < 1000$ km) about the point. Each point therefore has a different analogue. In effect, the AF procedure is like NWP; however, the "solutions" to the equations describing the atmosphere are not solved but determined from an empirical database of analogues. Theoretically, this concept can be extended to other meteorological variables where similarity of analogues could be assumed. The AF approach requires, of course, building a library of analogues applicable to the region. While interesting, FA would not be useful in practice for several years.

Ochs and Kidder (1989) developed a short-range forecasting system that was used in the field where the location had power, communications, and computer hardware. Their system, which advects GOES images with analyzed winds, provided a forecasting capability by integrating up-to-date analyzed products.

Recently, Brooks et al. (1992) discussed the difficulty of using mesoscale models in operational forecasting. They noted that the technology for the mesoscale, short-range forecast problem is far ahead of the theoretical development of the solution today. That is, the required observing, communications, and computer systems are available now; however, the conceptual framework and theoretical foundation for using them are not yet in place. Interestingly, the situation is in marked contrast to the dawn of the NWP age, when the theory of large-scale motions was becoming unified, but the technology was not ready to apply the theory.

2.5.4 Recommendation

As might be expected, there is a wide range of approaches to extending an analysis to a short-term forecast. This review has considered forecasts as simple as persistence, which is nearly useless during significant weather, and as rigorous as forecasts from sophisticated mesoscale NWP models, which have had only limited success so far. The most useful NWP models, however, were focused on forecasting specific weather events, such as heavy rain or severe weather. Such models require the assimilation and extrapolation of weather radar data and satellite imagery in addition to the conventional weather elements that RAP analyzes.

In between persistence and dynamic models are statistical techniques, which have several advantages over the other two methods. Statistical methods can be applied quickly at forecast time. They make use of the analysis as is, that is, without initialization, and are developed with available data. Since the statistical relationships are developed locally, they apply to the region of interest. Local effects are therefore accounted for indirectly without special consideration. Accordingly, statistical methods have greater appeal than dynamical methods for extending a RAP analysis.

The recommended solution is to test the technique discussed in the RAP proposal and reviewed in the background section. That is, the RWM forecast, valid 12 hr beyond the RAP analysis, should be modified by using the zero hour difference (analysis - first guess) field adjusted for the 12-hr time change. This solution is obviously the least costly of the viable solutions examined and has a high probability of success. The resulting short-term forecast can be compared to the RWM forecast for validation (or in OSSEs the short-term forecasts could be compared to truth). If successful, the RAP pseudoanalysis can then be compared to more sophisticated forecasts. Nevertheless, the requirements for specified forecast accuracy should determine whether the pseudoanalysis is acceptable.

3. SUMMARY AND CONCLUSIONS

The development of the RAP was both a technically challenging and rewarding project. The analysis procedure is unique in several respects. The multivariate forecast error correlation models, essential to the successful implementation of a regional optimum interpolation scheme, were developed by fitting real data to theoretically derived models. These models were used in conjunction with the FSR to perform a rigorous observation quality control of the entire set of observations available to an analysis, and then to select the "best" observations to perform an analysis at a point.

The resulting analyses were verified in real data and simulation modes. The experiments with real data showed that RAP analyses were a substantial improvement over the first guess forecast and analyses produced by HIRAS. In addition, the forecast error correlation models were shown to be better than physical models that did not rely on data fitted to them. The observing system simulation experiments provided a means for assessing the confidence limits of RAP analyses.

In conclusion, the RAP is a sound concept and a valuable prototype. Improved forecast error correlation models and more efficient numerical schemes could convert RAP into an operational regional analysis model.

REFERENCES

- Beran and MacDonald, 1982: *Nowcasting*. Academic Press, 256 pp.
- Bergman, K. H., 1979: Multivariate analysis of temperature and winds using optimum interpolation. *Mon. Wea. Rev.*, **107**, 1423-1443.
- Brooks, H. E., C. A. Doswell, III, and R. A. Maddox, 1992: On the use of mesoscale and cloud-scale models in operational forecasting. *Wea. Forecasting*, **7**, 120-131.
- Browning, K. A., 1982: *Nowcasting*. Academic Press, 256 pp.
- Buell, C. E., 1972: Correlation functions for winds and geopotential on isobaric surfaces. *J. Appl. Meteor.*, **11**, 51 -59.
- Burgeson, J., C. Redder, I. Halberstam, and R. Shapiro, 1991: Regional-Scale Analysis and Forecasting (RAP) Report. Interim Report, PL-TR-91-2135, ADA248384, Science and Technology Corp., Hampton, Virginia.
- Daley, R., 1985: The analysis of synoptic scale divergence by statistical interpolation procedure. *Mon. Wea. Rev.*, **107**, 1423-1443.
- Hollingsworth, A., and P. Lonnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38**, 111-136.
- Keegan, T. J., and R. Shapiro, 1985: *Methods of Meteorological Data Specification and Objective Field Analysis: A Review*. Technical Report AFGL-TR-85-0223, ADA160397, Air Force Geophysics Laboratory Atmospheric Sciences Division (AFGL/LYS), Hanscom Air Force Base, Massachusetts.

- Lonnerberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38**, 137-161.
- Mc Arthur, R. C., J. R. Davis, and D. Reynolds, 1987: Scenario-driven automatic pattern recognition in nowcasting. *J. Atmos. and Ocean Tech.*, **4**, 29-35.
- Mills, G. A., and R. S. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area mesoscale model. *Mon. Wea. Rev.*, **118**, 1250-1277.
- Mitchell, H. L., C. Charette, C. Chouinard, and B. Braslett, 1990: Revised interpolation statistics for the Canadian data assimilation procedure: Their derivation and application. *Mon. Wea. Rev.*, **118**, 1591-1614.
- Ochs, H. T., and S. Q. Kidder, 1989: Forecasting/nowcasting system for remote field locations. *J. Atmos. and Ocean Tech.*, **6**, 218-221.
- Petersen, R. A., and R. Homan, 1989: Short-range forecasting and nowcasting using a simple isentropic prediction model. *Wea. Forecasting*, **4**, 5-23.
- Pielke, R. A., 1981: Mesoscale numerical modeling. In *Advances in Geophysics*, **23**, 185-344, Academic Press.
- Science and Technology Corporation, RAP Initial Work Plan 1989: STC Technical Report 3072, Science and Technology Corp., Hampton, Virginia.
- Thiebaut, H. J., H. J. Mitchell, and D. W. Shantz, 1986: Horizontal structure of hemispheric forecast error correlations for geopotential and temperature. *Mon. Wea. Rev.*, **114**, 1048-1066.
- Thompson, P.D., 1988: Stochastic-dynamic prediction of the three-dimensional quasi-geostrophic flow. *J. Atmos. Sci.*, **45**, 2269-2279.

- Van den Dool, H. M., 1989: New look at weather through analogues. *Mon. Wea. Rev.*, **117**, 2230-2247.
- Warner, T. T., and N. L. Seaman, 1990: A real-time mesoscale numerical weather prediction system used for research, teaching and public service at the Pennsylvania State University. *Bull. Amer. Meteor. Soc.*, **71**, 792-805.
- Wilks, D. S., 1991: Representing serial correlations of meteorological events and forecasts in dynamic decision-analytic models. *Mon. Wea. Rev.*, **119**, 1640-1662.

APPENDIX A

CASE STUDIES OF THE BUDDY CHECK

Tables 1-10 illustrate the effect of both the buddy check and the gross error check on the analysis using a 12-hr first guess forecast and rawinsonde observations valid 1 January 1989, 12 UTC over central American and Eurasia. A visual inspection shows that the buddy and gross error checks significantly reduced the root-mean-square (rms) of the analysis residuals ϵ_a^\wedge (or errors). In some cases, the root-mean-square of ϵ_a^\wedge initially exceeded but finally lowered to a value significantly less than the root-mean-square of the forecast residual.

Table 11 contains a summary of the results in Tables 1-10 and consists of the percentage of the observations removed and the reduction of the root-mean-square of ϵ_a^\wedge . The expression for the percent of removal or reduction of a value is

$$\frac{|X_i - X_f|}{X_i} \times 100\% \quad (1)$$

where X_i and X_f are the initial and final values, respectively, and the value X in this study is either the number of observations or the root-mean-square of ϵ_a^\wedge . The percentages are averages over the pressure levels, 850 mb, 500 mb, 300 mb, 150 mb, and 70 mb. Note the higher percentage rate of observation removal among wind values than scalar values, which is a result of rejecting both wind components even if only one component fails either the buddy or gross error check.

Table 1a. A Comparison of the Effects on the Root-Mean-Square Height (m) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Eurasia

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	173	67.75	162	16.79	143	11.81
500 mb	161	52.31	159	24.99	140	16.15
300 mb	158	64.81	153	40.08	140	25.14
150 mb	145	70.37	142	45.17	128	32.53
70 mb	76	90.06	76	90.06	71	52.81

Table 1b. Same as Table 1a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	173	60.36	162	31.00	143	30.08
500 mb	161	56.65	159	37.95	140	35.56
300 mb	158	71.00	153	51.70	140	48.90
150 mb	145	79.56	142	62.48	128	57.64
70 mb	76	84.10	76	84.10	71	73.56

Table 2a. A Comparison of the Effects on the Root-Mean-Square Temperature (K) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Eurasia

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	163	2.004	161	1.948	149	1.471
500 mb	164	2.684	159	1.679	145	1.399
300 mb	159	2.226	151	1.646	133	.9922
150 mb	146	1.748	146	1.748	128	1.213
70 mb	75	2.753	71	2.312	66	1.267

Table 2b. Same as Table 2a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	163	2.112	161	1.979	149	1.876
500 mb	164	2.716	159	1.675	145	1.565
300 mb	159	1.990	151	1.506	133	1.310
150 mb	146	3.297	146	3.297	128	3.168
70 mb	75	4.616	71	3.195	66	3.101

Table 3a. A Comparison of the Effects on the Root-Mean-Square Dewpoint (K) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Eurasia (NA signifies data not available since RAP does not perform dewpoint analyses for levels above 300 mb)

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	159	4.023	156	3.810	141	2.678
500 mb	149	5.333	142	4.813	127	3.711
300 mb	86	4.889	82	4.422	76	3.525
150 mb	NA	NA	NA	NA	NA	NA
70 mb	NA	NA	NA	NA	NA	NA

Table 3b. Same as Table 3a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	159	3.995	156	3.458	141	2.908
500 mb	149	5.815	142	4.786	127	4.334
300 mb	86	5.667	82	4.479	76	3.879
150 mb	NA	NA	NA	NA	NA	NA
70 mb	NA	NA	NA	NA	NA	NA

Table 4a. A Comparison of the Effects on the Root-Mean-Square U-Wind (m/s) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Eurasia

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	176	4.546	170	3.442	151	2.778
500 mb	163	4.818	156	4.279	131	3.158
300 mb	153	5.718	143	4.563	117	3.502
150 mb	143	5.716	140	4.449	111	3.005
70 mb	67	5.075	67	5.075	57	3.002

Table 4b. Same as Table 4a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	176	4.231	170	3.575	151	3.450
500 mb	163	5.294	156	4.222	131	4.094
300 mb	153	7.033	143	5.221	117	4.912
150 mb	143	6.970	140	6.019	111	5.426
70 mb	67	8.519	67	8.519	57	8.065

Table 5a. A Comparison of the Effects on the Root-Mean-Square V-Wind (m/s) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Eurasia

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	176	3.895	170	3.652	151	2.878
500 mb	163	4.931	156	4.884	131	3.284
300 mb	153	5.833	143	5.877	117	3.609
150 mb	143	5.448	140	5.286	111	3.068
70 mb	67	6.588	67	6.588	57	4.771

Table 5b. Same as Table 5a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	176	4.163	170	3.951	151	3.806
500 mb	163	6.191	156	6.222	131	6.109
300 mb	153	8.068	143	8.234	117	7.941
150 mb	143	7.904	140	7.474	111	5.523
70 mb	67	6.666	67	6.666	57	6.690

Table 6a. A Comparison of the Effects on the Root-Mean-Square Height (m) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Central America

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	64	8.294	64	8.294	59	6.300
500 mb	64	21.56	63	16.10	59	11.71
300 mb	64	24.85	64	24.85	62	21.29
150 mb	63	37.04	63	37.04	57	22.43
70 mb	53	64.88	52	49.91	49	34.74

Table 6b. Same as Table 6a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	64	14.99	64	14.99	59	14.17
500 mb	64	23.32	63	19.79	59	18.68
300 mb	64	32.15	64	32.15	62	31.98
150 mb	63	50.17	63	50.17	57	44.14
70 mb	53	46.79	52	37.67	49	35.84

Table 7a. A Comparison of the Effects on the Root-Mean-Square Temperature (K) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Central America

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	61	1.672	60	1.714	57	1.456
500 mb	64	1.654	63	1.253	56	0.8980
300 mb	64	1.397	63	1.147	62	1.124
150 mb	63	1.728	63	1.728	60	1.232
70 mb	52	2.213	51	2.043	45	1.096

Table 7b. Same as Table 7a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	61	1.932	60	1.764	57	1.660
500 mb	64	1.792	63	1.472	56	1.377
300 mb	64	1.500	63	1.364	62	1.365
150 mb	63	3.042	63	3.042	60	3.104
70 mb	52	2.786	51	2.482	45	2.224

Table 8a. A Comparison of the Effects on the Root-Mean-Square Dewpoint (K) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Central America (NA signifies data not available because RAP does not perform dewpoint analyses above 300 mb)

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	61	7.073	56	3.862	52	3.521
500 mb	64	9.790	63	9.098	56	8.233
300 mb	30	11.67	22	4.804	20	2.021
150 mb	NA	NA	NA	NA	NA	NA
70 mb	NA	NA	NA	NA	NA	NA

Table 8b. Same as Table 8a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	61	6.380	56	4.340	52	4.246
500 mb	64	10.66	63	10.02	56	9.211
300 mb	30	9.324	22	3.753	20	3.297
150 mb	NA	NA	NA	NA	NA	NA
70 mb	NA	NA	NA	NA	NA	NA

Table 9a. A Comparison of the Effects on the Root-Mean-Square U-Wind (m/s) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Central America

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	61	1.993	61	1.993	50	1.823
500 mb	64	2.268	64	2.268	58	1.781
300 mb	62	3.633	62	3.633	52	3.064
150 mb	54	5.139	54	5.139	46	3.232
70 mb	45	3.380	45	3.380	39	2.138

Table 9b. Same as Table 9a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	61	3.467	61	3.467	50	3.379
500 mb	64	12.27	64	12.27	58	12.25
300 mb	62	20.50	62	20.50	52	19.75
150 mb	54	19.68	54	19.68	46	19.19
70 mb	45	10.20	45	10.20	39	9.742

Table 10a. A Comparison of the Effects on the Root-Mean-Square V-Wind (m/s) Analysis Residuals (errors) after Performing No Checks, Gross Error Checks, and Buddy Checks (with gross error checks) on the Set of Rawinsonde Observations (Obs) Valid 1 January 1989, 12 UTC over Central America

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	Analysis Error	Number of Obs	Analysis Error	Number of Obs	Analysis Error
850 mb	61	2.236	61	2.236	50	1.261
500 mb	64	1.744	64	1.744	58	1.577
300 mb	62	2.748	62	2.748	52	2.063
150 mb	54	2.709	54	2.709	46	1.998
70 mb	45	1.925	45	1.925	39	1.802

Table 10b. Same as Table 10a Except for Comparison of the Effects on the 12-Hr First Guess (FG) Forecast Residuals

Pressure Level	Not Checked		Gross Error Checked		Buddy Checked	
	Number of Obs	FG Error	Number of Obs	FG Error	Number of Obs	FG Error
850 mb	61	3.668	61	3.668	50	3.283
500 mb	62	2.540	64	2.540	58	2.608
300 mb	62	5.858	62	5.858	52	5.646
150 mb	54	6.118	54	6.118	46	5.928
70 mb	45	2.450	45	2.450	39	2.281

Table 11a. The Percent of the Root-Mean-Square Analysis Residual (error) Reduced During the Gross Error Check, Buddy Check, and the Combined Buddy and Gross Error Check on Rawinsonde Observations Valid 1 January 1989, 12 UTC over Eurasia (Also shown is the simultaneous percent removal of observations [Obs]. All values are means among the five pressure levels: 850 mb, 500 mb, 300 mb, 150 mb, and 70 mb.)

Variable	Gross Error Check		Buddy Check		Buddy and Gross Error Check	
	% Obs Removed	% Error Reduction	% Obs Removed	% Error Reduction	% Obs Removed	% Error Reduction
Z	2	40	10	34	12	61
T	3	16	9	33	12	43
T _d	4	5	9	24	13	31
u	3	15	16	28	18	43
v	3	2	16	33	18	33

Table 11b. Same as Table 11a Except for Observations Valid over Central America

Variable	Gross Error Check		Buddy Check		Buddy and Gross Error Check	
	% Obs Removal	% Error Reduction	% Obs Removal	% Error Reduction	% Obs Removal	% Error Reduction
Z	1	10	7	28	7	35
T	2	10	7	26	8	31
T _d	12	34	9	26	17	49
u	0	0	15	24	15	24
v	0	0	15	22	15	22

APPENDIX B

RESULTS OF OBSERVING SYSTEM SIMULATION EXPERIMENTS AT 850 AND 300 mb

Table 1. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Eurasia on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.508	8.671	NA	NA	3.306	26.85
24 hr BASELINE	3.596	17.13	NA	NA	4.963	39.63
36 hr BASELINE	3.884	30.05	NA	NA	4.743	45.61
12 hr CONTROL	5.755	10.72	6.020	50.51	5.269	27.80
24 hr CONTROL	5.963	20.20	6.795	49.38	6.889	40.00
36 hr CONTROL	6.645	33.60	7.119	44.69	6.438	46.21
12 hr TACOBS	24.00	20.53	19.27	57.44	14.23	34.98
24 hr TACOBS	22.78	24.46	12.79	53.20	24.58	43.06
36 hr TACOBS	22.70	36.43	15.67	45.12	20.85	47.22

Table 2. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.6449	1.806	NA	NA	0.6044	1.902
24 hr BASELINE	0.6008	1.862	NA	NA	0.6456	1.860
36 hr BASELINE	0.4318	2.012	NA	NA	0.6482	2.154
12 hr CONTROL	1.529	1.797	1.077	2.148	1.037	1.793
24 hr CONTROL	1.369	1.852	1.310	1.975	1.547	1.934
36 hr CONTROL	1.123	2.092	1.409	2.482	1.459	2.288
12 hr TACOBS	1.879	2.319	1.330	2.220	1.465	1.723
24 hr TACOBS	1.910	2.530	1.743	2.182	1.978	2.170
36 hr TACOBS	2.065	2.841	1.797	2.508	2.150	2.464

Table 3. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.011	5.330	NA	NA	2.465	3.837
24 hr BASELINE	2.772	5.075	NA	NA	3.039	4.217
36 hr BASELINE	2.203	4.899	NA	NA	3.464	5.228
12 hr CONTROL	4.739	5.286	2.678	3.892	3.792	4.352
24 hr CONTROL	3.899	5.326	4.053	4.693	4.959	4.759
36 hr CONTROL	4.464	5.213	4.715	5.398	5.204	5.461
12 hr TACOBS	6.500	6.451	4.930	5.032	5.036	5.046
24 hr TACOBS	5.904	5.912	5.396	5.576	5.395	5.644
36 hr TACOBS	4.760	5.316	5.600	5.690	5.764	6.018

Table 4. Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.8276	1.747	NA	NA	1.082	2.198
24 hr BASELINE	0.8376	2.076	NA	NA	1.441	2.686
36 hr BASELINE	1.209	2.930	NA	NA	1.718	3.932
12 hr CONTROL	2.512	2.009	2.797	2.292	2.351	2.458
24 hr CONTROL	2.692	2.309	2.409	3.011	2.527	2.795
36 hr CONTROL	2.787	3.188	2.650	3.469	2.446	4.214
12 hr TACOBS	2.767	2.990	3.298	2.672	3.240	3.540
24 hr TACOBS	3.172	3.121	2.802	2.927	2.967	3.530
36 hr TACOBS	3.287	3.722	3.175	3.416	3.198	4.484

Table 5. Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.9002	1.939	NA	NA	0.9261	2.190
24 hr BASELINE	0.8864	2.589	NA	NA	1.536	3.136
36 hr BASELINE	0.8644	3.356	NA	NA	1.465	4.242
12 hr CONTROL	2.676	2.159	2.580	2.881	2.184	2.531
24 hr CONTROL	2.496	2.876	2.565	3.381	2.796	3.471
36 hr CONTROL	2.676	3.675	2.877	3.838	2.512	4.220
12 hr TACOBS	3.733	4.262	2.905	2.936	3.021	3.136
24 hr TACOBS	3.559	4.447	3.172	3.534	3.555	4.073
36 hr TACOBS	3.683	4.889	2.642	3.984	4.244	4.605

Table 6. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	6.329	19.00	5.030	21.27	5.179	13.09
24 hr BASELINE	5.526	19.22	8.538	19.69	4.417	7.230
36 hr BASELINE	9.456	25.69	8.440	21.78	4.787	13.12
12 hr CONTROL	7.541	18.90	6.646	23.17	7.161	13.62
24 hr CONTROL	7.954	19.86	9.792	20.88	8.382	9.953
36 hr CONTROL	11.15	26.04	9.714	23.68	8.068	12.50
12 hr TACOBS	8.207	19.82	6.785	22.32	8.769	14.39
24 hr TACOBS	7.571	20.24	12.82	19.69	10.67	9.769
36 hr TACOBS	10.72	26.43	10.40	22.80	14.95	12.78

Table 7. **Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)**

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.054	2.514	1.148	2.387	0.7115	2.132
24 hr BASELINE	1.559	3.057	1.349	2.828	1.317	2.880
36 hr BASELINE	1.008	2.161	0.9860	2.172	0.8969	2.117
12 hr CONTROL	1.492	2.503	1.494	2.238	1.011	2.050
24 hr CONTROL	1.689	2.963	1.422	2.734	1.516	2.803
36 hr CONTROL	1.219	2.069	1.407	2.067	1.197	2.230
12 hr TACOBS	1.531	2.560	1.522	2.273	1.114	2.075
24 hr TACOBS	1.960	2.976	1.694	2.837	1.744	2.840
36 hr TACOBS	1.382	2.047	1.924	2.182	1.295	2.384

Table 8. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	4.896	5.177	8.593	9.081	7.333	7.344
24 hr BASELINE	6.239	6.540	7.438	9.562	5.771	7.003
36 hr BASELINE	7.467	9.209	7.520	10.97	5.764	7.226
12 hr CONTROL	6.322	5.539	9.000	8.452	11.07	6.968
24 hr CONTROL	6.553	7.091	7.740	8.426	7.304	6.485
36 hr CONTROL	8.100	9.621	7.744	10.27	6.959	7.179
12 hr TACOBS	6.314	5.811	9.556	9.389	11.63	7.385
24 hr TACOBS	7.033	7.232	8.309	8.955	7.809	6.668
36 hr TACOBS	8.034	9.653	7.778	10.33	6.748	6.806

Table 9. Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.678	2.845	2.216	3.188	1.874	2.081
24 hr BASELINE	3.174	3.264	2.564	3.773	2.415	3.275
36 hr BASELINE	2.067	3.081	2.359	3.835	2.074	3.883
12 hr CONTROL	3.906	3.262	3.347	3.846	3.034	2.344
24 hr CONTROL	3.798	3.428	4.272	4.011	3.181	3.089
36 hr CONTROL	3.262	3.163	3.750	3.708	2.948	3.557
12 hr TACOBS	3.942	3.078	3.560	4.031	3.377	2.578
24 hr TACOBS	4.326	3.577	3.632	4.234	3.830	3.314
36 hr TACOBS	3.623	3.537	4.130	3.824	3.251	3.808

Table 10. **Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)**

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.273	3.148	2.126	3.363	1.833	2.539
24 hr BASELINE	2.666	3.618	1.788	3.709	2.454	2.949
36 hr BASELINE	2.179	3.509	2.075	3.527	2.350	3.500
12 hr CONTROL	3.279	3.267	3.528	3.807	2.905	2.742
24 hr CONTROL	3.722	3.911	2.983	3.607	3.547	2.994
36 hr CONTROL	3.413	3.822	3.252	3.429	3.872	3.194
12 hr TACOBS	3.684	3.596	3.180	3.601	3.265	2.834
24 hr TACOBS	3.363	3.918	2.954	3.432	3.862	3.171
36 hr TACOBS	3.657	4.018	3.174	3.193	4.784	3.306

Table 11. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	6.184	19.00	5.115	21.27	5.087	13.09
24 hr BASELINE	5.526	19.22	8.471	19.69	4.417	7.230
36 hr BASELINE	6.455	25.69	8.452	21.78	4.777	13.12
12 hr CONTROL	7.586	18.90	6.617	23.17	7.181	13.62
24 hr CONTROL	7.954	19.86	10.00	20.88	8.382	9.953
36 hr CONTROL	8.591	26.07	9.726	23.68	8.249	12.50
12 hr TACOBS	8.777	19.82	6.916	22.32	8.792	14.39
24 hr TACOBS	7.571	20.24	11.72	19.69	10.67	9.769
36 hr TACOBS	10.17	26.43	10.25	22.80	12.60	12.78

Table 12. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.044	2.514	1.148	2.387	0.6876	2.132
24 hr BASELINE	1.562	3.057	1.356	2.828	1.317	2.880
36 hr BASELINE	0.9740	2.161	0.9866	2.172	0.9108	2.117
12 hr CONTROL	1.496	2.503	1.494	2.238	1.016	2.050
24 hr CONTROL	1.693	2.963	1.423	2.734	1.516	2.803
36 hr CONTROL	1.183	2.069	1.407	2.067	1.198	2.230
12 hr TACOBS	1.666	2.560	1.518	2.273	1.094	2.075
24 hr TACOBS	2.014	2.976	1.748	2.837	1.744	2.840
36 hr TACOBS	1.402	2.047	1.631	2.182	1.290	2.384

Table 13. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	4.896	5.177	8.593	9.081	7.333	7.344
24 hr BASELINE	6.239	6.540	7.438	9.562	5.771	7.003
36 hr BASELINE	7.467	9.209	7.520	10.97	5.764	7.226
12 hr CONTROL	6.322	5.539	9.000	8.452	11.07	6.968
24 hr CONTROL	6.553	7.091	7.740	8.426	7.304	6.485
36 hr CONTROL	8.100	9.621	7.744	10.27	6.959	7.179
12 hr TACOBS	6.314	5.811	9.556	9.389	11.63	7.385
24 hr TACOBS	7.033	7.232	8.309	8.955	7.809	6.668
36 hr TACOBS	8.034	9.653	7.778	10.33	6.748	6.806

Table 14. Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.678	2.845	2.216	3.188	1.874	2.081
24 hr BASELINE	3.175	3.264	2.546	3.773	2.415	3.275
36 hr BASELINE	2.073	3.081	2.360	3.835	2.074	3.883
12 hr CONTROL	3.906	3.262	3.347	3.846	3.038	2.344
24 hr CONTROL	3.798	3.428	4.266	4.011	3.181	3.089
36 hr CONTROL	3.262	3.163	3.750	3.708	2.949	3.557
12 hr TACOBS	3.942	3.078	3.560	4.031	3.377	2.578
24 hr TACOBS	4.326	3.577	3.630	4.234	3.830	3.314
36 hr TACOBS	3.610	3.537	4.129	3.824	3.294	3.808

Table 15. Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 850-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	2.273	3.147	2.126	3.363	1.833	2.539
24 hr BASELINE	2.662	3.618	1.790	3.709	2.454	2.949
36 hr BASELINE	2.179	3.509	2.075	3.527	2.350	3.500
12 hr CONTROL	3.279	3.267	3.528	3.807	2.905	2.742
24 hr CONTROL	3.718	3.911	2.988	3.607	3.547	2.994
36 hr CONTROL	3.413	3.822	3.252	3.429	3.872	3.194
12 hr TACOBS	3.684	3.596	3.180	3.601	3.265	2.834
24 hr TACOBS	3.356	3.918	2.894	3.432	3.862	3.171
36 hr TACOBS	3.684	4.018	3.169	3.193	4.796	3.306

Table 16. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Eurasia on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	10.16	16.31	NA	NA	9.758	30.66
24 hr BASELINE	10.79	28.78	NA	NA	10.77	32.86
36 hr BASELINE	11.47	51.96	NA	NA	11.63	41.40
12 hr CONTROL	16.82	17.89	15.41	58.24	16.11	32.27
24 hr CONTROL	15.80	29.90	14.31	45.07	16.38	32.41
36 hr CONTROL	20.11	53.98	17.59	49.78	18.80	42.69
12 hr TACOBS	30.51	31.08	22.22	61.36	38.67	42.45
24 hr TACOBS	34.55	50.03	26.38	50.90	39.54	45.16
36 hr TACOBS	50.17	71.44	39.73	53.85	45.85	51.46

Table 17. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.2711	0.8453	NA	NA	0.2424	0.5354
24 hr BASELINE	0.3168	0.7622	NA	NA	0.3539	0.7574
36 hr BASELINE	0.2690	0.8932	NA	NA	0.4807	1.270
12 hr CONTROL	1.460	0.9974	1.392	0.8308	1.025	0.7729
24 hr CONTROL	1.236	0.9114	1.284	0.9132	1.520	1.011
36 hr CONTROL	1.110	0.9790	1.177	1.398	1.556	1.465
12 hr TACOBS	1.545	1.022	1.544	0.9442	1.918	1.945
24 hr TACOBS	1.183	0.9646	1.281	1.016	2.477	2.296
36 hr TACOBS	1.450	1.254	1.426	1.520	2.954	2.570

Table 18. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Eurasia on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.558	2.022	NA	NA	1.612	2.657
24 hr BASELINE	1.309	2.231	NA	NA	1.472	3.157
36 hr BASELINE	1.679	3.093	NA	NA	1.624	4.028
12 hr CONTROL	2.679	2.533	2.869	1.937	3.968	3.249
24 hr CONTROL	2.747	2.788	2.677	2.595	3.195	3.391
36 hr CONTROL	3.345	3.659	3.751	3.767	3.155	4.118
12 hr TACOBS	2.745	3.323	3.773	2.486	5.688	3.813
24 hr TACOBS	3.948	3.530	3.139	2.689	4.260	3.956
36 hr TACOBS	3.582	3.795	4.441	3.907	4.288	4.095

Table 19. Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.177	2.673	NA	NA	0.9057	2.259
24 hr BASELINE	1.565	4.061	NA	NA	1.621	4.134
36 hr BASELINE	1.953	6.650	NA	NA	1.951	5.345
12 hr CONTROL	6.345	3.216	6.001	2.938	5.075	3.040
24 hr CONTROL	5.298	4.591	5.521	4.429	5.337	4.503
36 hr CONTROL	5.610	6.934	5.992	5.679	5.863	5.892
12 hr TACOBS	6.549	5.194	4.487	3.988	5.733	4.773
24 hr TACOBS	6.853	7.153	6.117	4.849	6.027	6.113
36 hr TACOBS	8.083	8.311	6.299	5.672	7.435	6.951

Table 20. Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Eurasia on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (All initial times at 00 UTC. An entry of NA means that the analysis was not available due to data processing problems.)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	1.113	2.657	NA	NA	0.9182	2.447
24 hr BASELINE	1.473	4.528	NA	NA	1.724	3.805
36 hr BASELINE	1.623	6.406	NA	NA	1.904	5.687
12 hr CONTROL	5.393	3.225	5.542	2.813	5.986	2.960
24 hr CONTROL	5.400	5.174	5.610	4.299	5.391	4.171
36 hr CONTROL	5.365	7.192	4.873	5.010	5.870	6.060
12 hr TACOBS	6.284	6.299	5.179	3.972	6.766	4.983
24 hr TACOBS	7.963	8.016	4.957	5.119	6.458	5.078
36 hr TACOBS	6.625	9.266	4.704	5.301	6.824	6.181

Table 21. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	14.65	33.93	13.79	39.41	15.22	34.47
24 hr BASELINE	19.97	41.37	18.12	48.36	15.77	38.12
36 hr BASELINE	18.82	48.89	19.99	62.73	13.96	29.82
12 hr CONTROL	22.45	35.49	23.68	39.75	22.97	34.98
24 hr CONTROL	21.42	40.38	19.12	51.18	21.05	39.42
36 hr CONTROL	20.96	53.01	23.11	65.67	18.36	30.97
12 hr TACOBS	24.72	36.92	26.75	39.11	25.03	35.82
24 hr TACOBS	27.67	41.98	20.60	47.73	24.68	36.58
36 hr TACOBS	27.13	53.60	24.10	64.72	20.53	29.99

Table 22. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.6235	0.8161	0.6115	1.098	0.5023	0.9529
24 hr BASELINE	0.8143	1.179	0.7651	1.138	0.5307	1.169
36 hr BASELINE	0.7589	1.334	0.5495	1.605	0.6225	0.9568
12 hr CONTROL	0.8872	0.8197	0.8448	1.261	0.8494	1.052
24 hr CONTROL	0.9824	1.397	0.9934	1.393	1.248	1.401
36 hr CONTROL	1.173	1.644	1.116	1.652	1.838	1.156
12 hr TACOBS	0.9960	1.035	0.8567	1.090	0.8498	0.9909
24 hr TACOBS	1.237	1.588	1.004	1.164	1.123	1.202
36 hr TACOBS	1.360	1.595	1.372	1.818	2.120	1.385

Table 23. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	4.532	5.262	4.975	4.795	3.903	4.737
24 hr BASELINE	5.635	5.572	6.218	5.484	5.610	5.688
36 hr BASELINE	4.714	6.870	4.822	5.136	3.876	6.039
12 hr CONTROL	10.47	6.336	16.61	5.441	10.63	5.013
24 hr CONTROL	13.35	6.565	15.28	5.816	9.962	5.679
36 hr CONTROL	9.123	7.749	7.647	6.130	10.17	6.278
12 hr TACOBS	14.51	5.914	17.03	5.478	12.19	5.613
24 hr TACOBS	21.82	6.073	9.524	6.080	10.80	6.085
36 hr TACOBS	8.521	7.101	10.00	6.311	10.17	7.020

Table 24. **Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)**

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.297	4.590	2.626	3.802	1.961	2.714
24 hr BASELINE	4.534	3.966	3.148	4.419	2.576	3.786
36 hr BASELINE	3.698	5.605	4.974	6.716	2.956	4.673
12 hr CONTROL	9.014	5.430	5.816	4.535	5.180	2.941
24 hr CONTROL	7.854	5.498	5.928	4.554	4.893	3.314
36 hr CONTROL	6.036	6.974	10.30	6.288	5.307	3.934
12 hr TACOBS	9.312	6.048	5.987	5.212	5.628	3.557
24 hr TACOBS	7.810	6.016	6.745	5.324	4.516	3.781
36 hr TACOBS	5.926	7.336	11.21	6.533	6.884	4.586

Table 25. Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases with Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.191	4.895	2.736	3.946	1.866	4.033
24 hr BASELINE	4.306	6.340	2.938	5.621	2.216	3.400
36 hr BASELINE	3.757	7.703	3.211	6.401	2.677	3.555
12 hr CONTROL	7.232	5.899	6.158	5.484	6.491	4.195
24 hr CONTROL	6.641	8.101	6.654	6.851	6.938	3.842
36 hr CONTROL	6.067	8.972	4.582	6.723	4.793	3.805
12 hr TACOBS	7.803	6.724	7.465	5.000	6.492	4.482
24 hr TACOBS	7.119	8.613	6.312	6.928	6.762	3.693
36 hr TACOBS	7.039	8.646	4.342	7.828	5.336	4.679

Table 26. Root-Mean-Square Height Errors (m) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	13.79	33.93	13.51	39.41	14.77	34.47
24 hr BASELINE	19.97	41.37	18.58	48.36	15.77	38.12
36 hr BASELINE	19.59	48.89	19.99	62.73	13.40	29.82
12 hr CONTROL	21.14	35.49	23.55	39.75	22.65	34.98
24 hr CONTROL	21.42	40.38	19.53	51.18	21.05	39.43
36 hr CONTROL	22.01	53.01	23.08	65.67	18.22	30.97
12 hr TACOBS	25.89	36.92	26.72	39.11	25.48	35.82
24 hr TACOBS	27.67	41.98	22.64	47.73	24.68	36.58
36 hr TACOBS	26.89	53.60	24.68	64.72	20.99	29.99

Table 27. Root-Mean-Square Temperature Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	0.5800	0.8161	0.6113	1.098	0.5002	0.9529
24 hr BASELINE	0.8019	1.179	0.7618	1.138	0.5307	1.169
36 hr BASELINE	0.7445	1.334	0.5526	1.605	0.6481	0.9568
12 hr CONTROL	0.8598	0.8197	0.8448	1.261	0.8508	1.052
24 hr CONTROL	0.9861	1.397	0.9937	1.393	1.248	1.401
36 hr CONTROL	1.166	1.644	1.119	1.652	1.825	1.156
12 hr TACOBS	0.8780	1.035	0.8577	1.090	0.8540	0.9909
24 hr TACOBS	1.263	1.588	0.9396	1.164	1.123	1.202
36 hr TACOBS	1.581	1.595	1.441	1.818	2.183	1.385

Table 28. Root-Mean-Square Dewpoint Errors (K) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	4.532	5.262	4.975	4.795	3.903	4.737
24 hr BASELINE	5.635	5.572	6.218	5.484	5.610	5.688
36 hr BASELINE	4.714	6.870	4.822	5.136	3.876	6.039
12 hr CONTROL	10.47	6.336	16.61	5.441	10.63	5.013
24 hr CONTROL	13.35	6.564	15.28	5.816	9.962	5.679
36 hr CONTROL	9.123	7.748	7.647	6.130	10.17	6.278
12 hr TACOBS	14.51	5.914	17.03	5.478	12.19	5.613
24 hr TACOBS	21.82	6.073	9.524	6.080	10.80	6.085
36 hr TACOBS	8.521	7.101	10.00	6.311	10.17	7.020

Table 29. Root-Mean-Square U-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.305	4.590	2.626	3.802	1.961	2.714
24 hr BASELINE	4.534	4.409	3.155	4.419	2.576	3.786
36 hr BASELINE	3.686	5.605	4.974	6.716	2.958	4.673
12 hr CONTROL	9.017	5.430	5.816	4.535	5.180	2.941
24 hr CONTROL	7.854	5.498	5.930	4.554	4.893	3.314
36 hr CONTROL	6.027	6.974	10.30	6.288	5.302	3.934
12 hr TACOBS	9.355	6.048	5.984	5.212	5.633	3.557
24 hr TACOBS	7.810	6.016	6.848	5.324	4.516	3.781
36 hr TACOBS	5.866	7.336	11.20	6.533	6.878	4.586

Table 30. **Root-Mean-Square V-Wind Errors (m/s) of RAP Analyses and First Guess Forecasts over Central America on the 300-mb Surface for the Given Forecast Periods and Observation Databases Without Satellite Data Included (all initial times at 00 UTC)**

	21 Jan Analysis	First Guess	26 Jan Analysis	First Guess	31 Jan Analysis	First Guess
12 hr BASELINE	3.160	4.895	2.736	3.946	1.866	4.033
24 hr BASELINE	4.306	6.340	2.938	5.621	2.216	3.400
36 hr BASELINE	3.796	7.703	3.212	6.401	2.664	3.555
12 hr CONTROL	7.215	5.899	6.158	5.484	6.491	4.195
24 hr CONTROL	6.641	8.101	6.654	6.851	6.938	3.842
36 hr CONTROL	6.156	8.972	4.583	6.723	4.791	3.805
12 hr TACOBS	7.770	6.724	7.474	5.000	6.492	4.482
24 hr TACOBS	7.119	8.613	6.247	6.928	6.762	3.693
36 hr TACOBS	7.064	8.646	4.325	7.828	5.309	4.679